



CENTRO INTERNACIONAL DE ESTUDOS  
DE DOUTORAMENTO E AVANZADOS  
DA USC (CIEDUS)

TESIS DE DOCTORADO

**EVALUACIÓN EMPÍRICA DE UN MODELO  
TEÓRICO DE LOS EFECTOS DEL FAKING  
SOBRE LAS MEDIDAS DE PERSONALIDAD  
OCUPACIONAL**

ALEXANDRA MARTÍNEZ GÓMEZ

**ESCUELA DE DOCTORADO INTERNACIONAL (EDI) EN ARTES Y  
HUMANIDADES, CIENCIAS SOCIALES Y JURÍDICAS**

**PROGRAMA DE DOCTORADO EN PSICOLOGÍA DEL TRABAJO Y LAS  
ORGANIZACIONES, JURÍDICA-FORENSE Y DEL CONSUMIDOR Y  
USUARIO**

SANTIAGO DE COMPOSTELA

2019





## **DECLARACIÓN DE LA AUTORA DE LA TESIS**

### **Evaluación empírica de un modelo teórico de los efectos del faking sobre las medidas de personalidad ocupacional**

Dña. Alexandra Martínez Gómez

Presento mi tesis, siguiendo el procedimiento adecuado al Reglamento, y declaro que:

- 1) La tesis abarca los resultados de la elaboración de mi trabajo.
- 2) En su caso, la tesis hace referencia a las colaboraciones que tuvo este trabajo.
- 3) La tesis es la versión definitiva presentada para su defensa y coincide con la versión enviada en formato electrónico.
- 4) Confirmo que esta tesis no incurre en ningún tipo de plagio de otros autores ni de trabajos presentados por mí para la obtención de otros títulos.

En Santiago de Compostela a 9 de octubre de 2019

Fdo.: Alexandra Martínez Gómez





## **AUTORIZACIÓN DE LOS DIRECTORES / TUTORES DE LA TESIS**

### **Evaluación empírica de un modelo teórico de los efectos del faking sobre las medidas de personalidad ocupacional**

D. Jesús Fernando Salgado Velo y Dña. Silvia Moscoso Ruibal

INFORMAN:

Que la presente tesis corresponde con el trabajo realizado por Dña. Alexandra Martínez Gómez bajo nuestra dirección, y autorizamos su presentación considerando que reúne los requisitos exigidos en el Reglamento de Estudios de Doctorado de la USC, y que como directores de ésta no incurrimos en causas de abstención establecidas en la Ley 40/2015.

Santiago de Compostela, 9 de octubre de 2019

Fdo: Jesús Fernando Salgado Velo

Fdo: Silvia Moscoso Ruibal



*Aos meus pais*







## AGRADECIMIENTOS

La realización de una tesis doctoral supone un importante reto y esfuerzo para cualquier persona que decida embarcarse en esta “aventura”. En mi caso esta oportunidad llegó inesperadamente hace cinco años y, sin duda, ha sido el mayor desafío al que me enfrentado hasta el momento. Hoy puedo decir que ha merecido la pena.

Finalizar esta tesis doctoral supone para mí la meta de una etapa de gran aprendizaje profesional y desarrollo personal. Son innumerables las experiencias que se han quedado guardadas en mi memoria y de las que me siento tremendamente orgullosa y agradecida por haberlas vivido. Pero no habría llegado hasta aquí sin el apoyo, dedicación y orientación de muchas personas que han sido partícipes de este proceso.

Me gustaría comenzar agradeciendo al Dr. Jesús F. Salgado y a la Dra. Silvia Moscoso por la oportunidad que me han brindado. Me siento muy orgullosa de formar parte de su equipo y estoy convencida de que no hubiera podido tener mejores directores de tesis. Su estímulo y guía han sido fundamentales en la finalización de esta tesis doctoral.

Quiero agradecer especialmente al Dr. Jesús F. Salgado por haber sido el primero en apostar por mí y darme la confianza como para creer que podría llegar hasta aquí. Estaré siempre en deuda con él por las horas que me ha prestado para resolver infinidad de dudas, por su sabiduría, paciencia, comprensión y apoyo y por sus palabras de ánimo en los momentos más duros, especialmente en estos últimos meses en los que el agobio y la ilusión libraban una batalla constante. Infinitas gracias también a la Dra. Silvia Moscoso, su ayuda y sus sabios consejos han sido fundamentales para lograr este objetivo. Su experiencia ha sido una fuente de estímulo para mi desarrollo profesional. Les estaré siempre agradecida por la confianza que han depositado en mi desde el primer momento.

Del mismo modo, quiero expresar mi gratitud al Dr. Marco Dipolo por la amabilidad y la hospitalidad con la que me ha acogido en su grupo de investigación. Por su ayuda y dedicación y por poner a mi disposición todos los medios y servicios para que

podiese sentirme como en casa durante mi estancia pre-doctoral en la Universidad de Bolonia. Muchísimas gracias también a Michela Vignoli por recibirme tan cálidamente, ha sido todo un placer conocerla y conocer de primera mano sus investigaciones.

Debo también agradecer al Ministerio de Economía y Competitividad por haber financiado la realización de esta tesis doctoral con la beca FPI que me fue otorgada, asociada al proyecto de investigación PSI2014-56615P concedido al Dr. Jesús F. Salgado.

Gracias a mis compañeras, que hoy puedo considerar amigas, Dámaris Cuadrado, Pamela Alonso e Inmaculada Otero, por las horas que han invertido en que esta tesis saliese adelante con su ayuda en la recogida de datos y en la revisión del texto. Pero sobre todo por su aliento, comprensión y generosidad. Sin ellas este camino hubiera sido mucho más duro. Quiero agradecer a Dámaris y a Pamela haberme acogido como una más desde el minuto uno. A Dámaris por su dedicación siempre que la he necesitado y por hacerme reír hasta en el día de mayor agobio. A Pamela por despertarme la ilusión para iniciar este “viaje” y por las innumerables dudas que ha tenido la paciencia de escuchar y ayudarme a resolver. Y a Inma, con la que inicié esta etapa, por haber permitido que compartamos nuestros miedos e inquietudes, pero también nuestras alegrías durante todo este tiempo. No podría pensar en una mejor compañera de “aventura”. Tampoco quiero olvidarme de nuestro compañero José Carlos Tavares, quiero expresarle todo mi cariño y admiración, su infinita sabiduría ha permitido que en cada conversación aprendiese algo nuevo.

Pero sin duda esta tesis es gracias al apoyo de personas que aun sin llegar a comprender el significado de este trabajo han permanecido a mi lado durante todo este tiempo y también a aquellas que, gracias al mismo, han aparecido para quedarse. Amigos, familia y compañeros que quizás sin saberlo han sido la energía que me ha mantenido a flote. Vane, Eli, Cris, Flavia, Vero, Laura, Lore, Lucía, gracias infinitas. Pero, especialmente, gracias a los dos pilares de mi vida. Félix, tu apoyo ha sido fundamental para seguir aquí. En los mejores, pero sobre todo en los peores momentos, has permanecido a mi lado y has sido capaz de sacar lo mejor de mí. Gracias por querer formar parte de mi vida. Y, por supuesto, a ti mamá. Gracias por apoyarme sin preguntar y por hacerme crecer como hija y como persona. Sin ti no hubiese conseguido alcanzar esta meta. Esto es gracias a ti.

## RESUMEN

El faking es un problema importante en la evaluación de la personalidad en selección personal y en los procesos de admisión de alumnos en instituciones académicas. La teoría psicométrica del faking (Salgado, 2016) ha demostrado los efectos del faking sobre los cuestionarios *single-stimulus* (SS): aumento de la media, disminución de la desviación típica y, en consecuencia, reducción de la fiabilidad, modificación de la estructura factorial y reducción de la validez de criterio. Sin embargo, tales efectos no se habían examinado en cuestionarios de elección forzosa (EF). Esta tesis pone a prueba las predicciones de esta teoría en los cuestionarios de EF, mediante la realización de tres estudios.

El primero consistió en un meta-análisis de los efectos del faking sobre las propiedades psicométricas de los cuestionarios de EF. Los resultados muestran que los tres formatos de EF, normativo, ipsativo y quasi-ipsativo, presentan resistencia al faking, aunque con diferente magnitud. Además, se ha encontrado que el tipo de diseño experimental afecta a los resultados.

El segundo estudio examinó la invarianza o equivalencia de la medida en los cuestionarios de EF quasi-ipsativos en condiciones de respuesta honesta y de faking. Los resultados indican que la estructura factorial de estos cuestionarios es robusta a los efectos del faking.

El tercer estudio investigó los efectos del faking sobre la validez predictiva de los cuestionarios de EF quasi-ipsativos. Los resultados señalan que el factor de conciencia es el mejor predictor del desempeño académico, incluso en condiciones de faking. También indican que la medida de desempeño es un moderador relevante de la validez predictiva.

Para finalizar se plantean las implicaciones teóricas y prácticas de los hallazgos y se proponen cuestiones de investigación para futuros estudios.

**Palabras clave:** *teoría psicométrica del faking; cuestionarios de elección forzosa; personalidad; invarianza de la medida; fiabilidad y validez*



## RESUMO

O faking é un problema importante na avaliación da personalidade para a selección de persoal e os procesos de admisión de estudantes en institucións académicas. A teoría psicométrica do faking (Salgado, 2016) demostrou os efectos do faking sobre os cuestionarios *single-stimulus* (SS): aumento da media, diminución da desviación estándar e, en consecuencia, redución da fiabilidade, modificación da estrutura factorial e redución da validez de criterio. Non obstante, tales efectos non se examinaron nos cuestionarios de elección forzosa (EF). Esta tese pon a proba as predicións desta teoría nos cuestionarios de EF, mediante a realización de tres estudos.

O primeiro consistiu nunha meta-análise dos efectos do faking sobre as propiedades psicométricas dos cuestionarios EF. Os resultados mostran que os tres formatos de EF, normativo, ipsativo e quasi-ipsativo, teñen resistencia ó faking, aínda que con distinta magnitude. Ademais, comprobouse que o tipo de deseño experimental afecta aos resultados.

O segundo estudo examinou a invarianza ou equivalencia da medida nos cuestionarios EF quasi-ipsativos en condicións de resposta honesta e de faking. Os resultados indican que a estrutura dos factores destes cuestionarios é robusta aos efectos do faking.

O terceiro estudo investigou os efectos do faking na validez de criterio dos cuestionarios EF quasi-ipsativos. Os resultados indican que o factor de conciencia é o mellor predictor do desempeño académico, incluso en condicións de faking. Tamén indican que a medida de desempeño é un moderador relevante de validez predictiva.

Finalmente, plantéxanse as implicacións teóricas e prácticas dos resultados e propóñense cuestións de investigación para futuros estudos.

**Palabras chave:** *teoría psicométrica do faking; cuestionarios de elección forzosa; personalidade; invarianza da medida; fiabilidade e validez*



# ABSTRACT

Faking is a critical problem for personality assessment in personnel selection procedures and academic admission processes. The psychometric theory of faking (Salgado, 2016) has shown that faking has a negative impact on *single stimulus* (SS) tests: overestimation of the average score, underestimation of the standard deviation, and consequently, modification of the factorial structure, and underestimation of the reliability and the predictive validity. However, these effects had not been yet examined in forced choice questionnaires (FC). By conducting three empirical studies, this dissertation tests the cited theory in FC.

The first study consisted of a meta-analysis of the effects of faking on the psychometric characteristics of FC. The results show that the three FC formats, normative, ipsative, and quasi-ipsative, present faking resistance, although to a different extent. Furthermore, the type of experimental design also affects the results.

The second study examined the measure invariance or equivalence in the quasi-ipsative FC tests under honest and faking conditions. The results show that the factorial structure of this type of FC is resistant to faking.

The third study examined the effects of faking on the predictive validity of the FC quasi-ipsative questionnaires. The results show that the dimension of personality conscientiousness is the better predictor of academic performance, even under faking conditions. They also prove that the type of performance measure is a moderating variable of the predictive validity.

Last, theoretical and practical implications are posited, and suggestions for future research are done.

**Keywords:** *psychometric theory of faking; forced choice questionnaires; personality; measure invarianza; reliability and validity*





## ÍNDICE

<b>INTRODUCCIÓN.....</b>	<b>27</b>
--------------------------	-----------

### **MARCO TEÓRICO**

#### **Aproximaciones Teóricas, Efectos, Medición y Control del Faking**

<b>CAPITULO 1 Faking: Concepto y Características.....</b>	<b>37</b>
1. FAKING: CONCEPTO .....	39
2. TAXONOMÍAS DEL FAKING.....	44
3. OPERATIVIZACIÓN DEL FAKING .....	47
3.1 Comparaciones Entre Grupos.....	48
3.2 Diferencias de Medias.....	50
3.3. Escalas de Deseabilidad Social .....	50
4. PREVALENCIA DEL FAKING.....	51
<b>CAPITULO 2 Personalidad y Faking.....</b>	<b>59</b>
1. PERSONALIDAD: MODELO DE LOS CINCO GRANDES .....	61
2. VALIDEZ PREDICTIVA DEL MODELO DE LOS CINCO GRANDES .....	64
2.1 Modelo de los Cinco Grandes Factores y Desempeño Ocupacional.....	65
2.2 Modelo de los Cinco Grandes Factores y Desempeño Académico .....	68
2.3 Modelo de los Cinco Grandes Factores y Validez Añadida .....	69
3. EFECTOS DEL FAKING SOBRE LAS MEDIDAS DE PERSONALIDAD.....	70
3.1 Efectos del Faking sobre las Puntuaciones .....	70
3.2 Efectos Sobre la Fiabilidad.....	77
3.3 Efectos del Faking Sobre la Validez .....	79
3.4 Efectos Sobre el Ranking de los Candidatos.....	88
<b>CAPITULO 3 Modelos Teóricos Explicativos del Faking.....</b>	<b>93</b>
1. VARIABLES ANTECEDENTES ANALIZADAS EN LA LITERATURA CIENTÍFICA .....	95
1.1 Características Individuales o Disposicionales.....	96
1.2 Características Situacionales o Contextuales.....	101
1.3 La Motivación Como Antecedente del Faking .....	103
2. TEORÍAS SOBRE LOS ANTECEDENTES DEL FAKING .....	103
2.1 Teoría de Snell et al. (1999).....	104
2.2 Teoría de McFarland y Ryan (2000) .....	106
2.3 Teoría de Mueller-Hanson et al. (2006) .....	108
2.4 Teoría de Goffin y Boyd (2009) .....	109

2.5 Teoría de la Clasificación Contrato-Rasgo (Griffith et al., 2011) .....	111
2.6 Teoría de Marcus (2009) .....	113
2.7 Modelo Autoregulatorio del Faking (Komar, Komar, Robie y Taggar, 2010) .....	115
2.8 Teoría de la Activación de los Rasgos (Tett y Simonet, 2011).....	115
2.9 Teoría de los Heurísticos (Kuncel, Goldberg y Kiger, 2011) .....	116
2.10 Teoría VIE: Valencia-Instrumentalidad-Expectativa (Ellingson y McFarland, 2011).....	117
3. TEORÍAS SOBRE LAS CONSECUENCIAS DEL FAKING .....	118
3.1 Teoría del Faking como Error Transitorio (Heggstad, George et al, 2006) .	118
3.2 Teoría del Faking como Error Sistemático de la Medida (Ziegler y Buehner, 2009).....	120
3.3 Modelo Teórico de los Efectos Psicométricos del Faking (Salgado, 2016) ....	121
<b>CAPITULO 4 Métodos de Detección y Control del Faking .....</b>	<b>127</b>
1. EVALUACIÓN DEL FAKING .....	129
2. MÉTODOS QUE DETECTAN EL FAKING .....	130
2.1 Escalas de Detección de Deseabilidad Social (EDS) .....	131
2.2. Técnica de la Respuesta Exagerada o Ítems Trampa (ítems bogus) .....	135
2.3. Baremos Basados en Muestras de Solicitantes .....	136
2.4 Tiempo de Latencia .....	137
2.5 Ítems Idiosincrásicos.....	139
2.6 Detección A Través del Lenguaje .....	139
2.7 Técnicas TRI.....	141
2.8 Métodos de Corrección del Faking.....	141
3. MÉTODOS DE CONTROL DEL FAKING .....	143
3.1 Estrategias de Advertencia .....	143
3.2 Test de Razonamiento Condicionado.....	145
<b>CAPITULO 5 Cuestionarios de Elección Forzosa y el Modelo Psicométrico de los Efectos del Faking .....</b>	<b>147</b>
1. LOS CUESTIONARIOS DE ELECCIÓN FORZOSA.....	149
1.1 Características de los Cuestionarios de EF.....	149
1.2 Propiedades Psicométricas de los Cuestionarios de EF .....	153
1.3 Validez Predictiva de los Cuestionarios de los Cuestionarios de EF.....	156
2. CUESTIONARIOS DE ELECCIÓN FORZOSA Y FAKING .....	158

## PARTE EMPIRICA

### Evaluación Empírica del Modelo Teórico en Cuestionarios de Elección Forzosa

<b>ESTUDIO 1 Meta-Análisis de los Efectos Psicométricos del Faking sobre los Cuestionarios de Personalidad de Elección Forzosa .....</b>	<b>167</b>
Introducción .....	169
Hipótesis de Estudio .....	171
Método .....	172
1. <i>Búsqueda de Estudios</i> .....	172
2. <i>Criterios de Inclusión</i> .....	173
3. <i>Codificación de Estudios</i> .....	175
4. <i>Reglas de Decisión</i> .....	175
5. <i>Análisis de Variables Moderadoras</i> .....	179
6. <i>Estimación del Tamaño del Efecto</i> .....	180
7. <i>Fiabilidad</i> .....	181
RESULTADOS .....	183
1. <i>Resultados de los Meta-Análisis Según el Diseño</i> .....	184
2. <i>Resultados de los Meta-análisis Jerárquicos Según el Tipo de Cuestionario de Elección Forzosa</i> .....	187
DISCUSIÓN .....	193
<i>Limitaciones del Estudio</i> .....	194
<i>Implicaciones Prácticas</i> .....	195
<i>Sugerencias de Investigación</i> .....	195
 <b>ESTUDIO 2 Efectos del Faking sobre la Estructura Factorial de los Cuestionarios de Personalidad de Elección Forzosa Quasi-Ipsativos: Invarianza/Equivalencia de la Medida .....</b>	<b>197</b>
INTRODUCCIÓN .....	199
MÉTODO .....	200
1. <i>Muestra</i> .....	200
2. <i>Instrumento de Medida de la Personalidad</i> .....	201
3. <i>Procedimiento</i> .....	203
RESULTADOS .....	204
1. <i>Análisis Factorial Exploratorio de la Estructura del QI5F_tri</i> .....	205
2. <i>Análisis de Componentes Principales de la Estructura del QI5F_tri</i> .....	221
3. <i>Análisis Factorial Confirmatorio de la Estructura del QI5F_tri</i> .....	230
DISCUSIÓN .....	232
<i>Implicaciones Prácticas</i> .....	234
<i>Limitaciones del Estudio</i> .....	234
 <b>ESTUDIO 3 Efectos del Faking sobre la Validez Predictiva de los Cuestionarios de Personalidad de Elección Forzosa Quasi-Ipsativos .....</b>	<b>237</b>
INTRODUCCIÓN .....	239
MÉTODO .....	242

1. Muestra .....	242
2. Instrumentos de medida .....	243
3. Procedimiento .....	244
RESULTADOS .....	246
DISCUSIÓN .....	264
<i>Implicaciones de los Resultados para la teoría y la práctica de la evaluación de la Personalidad</i> .....	267
<i>Limitaciones del estudio</i> .....	268
<b>CONCLUSIONES GENERALES .....</b>	<b>271</b>
<i>Impliaciones para la Práctica</i> .....	277
<i>Sugerencias para Investigaciones Futuras</i> .....	278
<i>Limitaciones</i> .....	278
<b>ANEXOS .....</b>	<b>281</b>
<b>REFERENCIAS .....</b>	<b>293</b>



## ÍNDICE DE TABLAS

Tabla 1. Conceptualización de los Tipos de Respuesta Distorsionada .....	42
Tabla 2. Revisión de Estudios sobre la Prevalencia del Faking de Griffith y Converse (2012).....	52
Tabla 3. Resultados del Meta-análisis de la Prevalencia del Faking.....	56
Tabla 4. Resultados del Meta-análisis de Hooper (2007) que Compara los Efectos del Faking entre Estudios Experimentales y Correlacionales para los Big Five.....	73
Tabla 5. Resultados del Meta-análisis de Salgado (2016) que Compara los Efectos del Faking entre Tres Tipos de Muestras: Candidatos, Empleados y Diseño Intra- grupo .....	74
Tabla 6. Coeficientes de Homogeneidad del Meta-análisis de Salgado (2016) entre Muestras de Trabajadores Motivados y no Motivados a Cometer Faking.....	83
Tabla 7. Coeficientes del Análisis Confirmatorio de la Estructura Factorial del Estudio de Cellar, Miller, Doverspike y Klawnsky (1996).....	87
Tabla 8. Características Individuales Relacionadas con la Tendencia a Cometer Faking .....	96
Tabla 9. Características Situacionales Relacionadas con la Tendencia a Cometer Faking.....	101
Tabla 10. Breve Descripción de las Escalas de Deseabilidad Social Incluidas en Cuestionarios de Personalidad .....	133
Tabla 11. Resultados del Meta-análisis de Nguyen y McDaniel (2000) sobre los Efectos del Faking en los Cuestionarios de EF según el Tipo de Diseño.....	160
Tabla 12. Resultados del Meta-análisis General de Cao y Drasgow (2019) de los efectos del Faking en los Cuestionarios de EF.....	161
Tabla 13. Cuestionarios de Personalidad Incluidos en el Meta-análisis .....	174
Tabla 14. Distribuciones de los Coeficientes de Fiabilidad de las Variables Incluidas en los Meta-análisis .....	182
Tabla 15. Meta-análisis de la Resistencia al Faking de los Cuestionarios de Elección Forzosa según el Tipo de Diseño .....	186
Tabla 16. Meta-análisis Jerárquico de la Resistencia al Faking de los Cuestionarios de Elección Forzosa Ipsativos.....	190
Tabla 17. Meta-análisis Jerárquico de la Resistencia al Faking de los Cuestionarios de Elección Forzosa Quasi-Ipsativos .....	191
Tabla 18. Meta-análisis de la Resistencia al Faking de los Cuestionarios de Elección Forzosa Normativos .....	192
Tabla 19. Características Demográficas de la Muestra ( $N = 1,141$ ).....	201
Tabla 20. Agrupación de los Ítems del Cuestionario de EF en Clústers .....	205
Tabla 21. Estadísticos Descriptivos de las Variables para la Muestra Total.....	206

Tabla 22. Matriz de Cargas Factorial Rotadas del Análisis Factorial Exploratorio del QI5F_tri Contestado Bajo Instrucciones de Honestidad, Empleando la Muestra Total ( $N = 939$ ).....	207
Tabla 23. Indicadores de Bondad de Ajuste del Modelo de Cinco Factores con Instrucciones de Respuesta Honesta al QI5F_Tri, Empleando la Muestra Total ( $N = 939$ ).....	208
Tabla 24. Matriz de Cargas Factoriales Rotadas del Análisis Factorial Exploratorio del QI5F_tri Contestado Bajo Instrucciones de Faking, Empleando la Muestra Total ( $N = 692$ ).....	209
Tabla 25. Indicadores de Bondad de Ajuste del Modelo de Cinco Factores con Instrucciones de Respuesta que Inducían Faking al QI5F/tri Empleando la Muestra Total ( $N = 692$ ) .....	210
Tabla 26. Estadísticos Descriptivos de las Variables en el Diseño Intra-Grupo ( $N = 490$ ).....	211
Tabla 27. Matriz de Cargas Factoriales Rotadas del Análisis Factorial Exploratorio del QI5F_tri Contestado Bajo Instrucciones de Honestidad, con un Diseño Intra-Grupo ( $N = 490$ ).....	212
Tabla 28. Indicadores de Bondad de Ajuste del Modelo de Cinco Factores con Instrucciones de Respuesta Honesta al QI5F_tri con un Diseño Intra-Grupo ( $N = 490$ ) .....	213
Tabla 29. Matriz de Cargas Factoriales Rotadas del Análisis Factorial Exploratorio del QI5F_tri Contestado Bajo Instrucciones de Faking, con un Diseño Intra-grupo ( $N = 490$ ).....	214
Tabla 30. Indicadores de Bondad de Ajuste del Modelo de Cinco Factores con Instrucciones de Respuesta que Inducían al QI5F_tri, con un Diseño Intra-Grupo ( $N = 490$ ).....	215
Tabla 31. Estadísticos Descriptivos de la Variables en el Diseño Inter-Grupos .....	216
Tabla 32. Matriz de Cargas Factoriales Rotadas del Análisis Factorial Exploratorio del QI5F_tri Contestado Bajo Instrucciones Honestas, con un Diseño Entre-Grupos ( $N = 449$ ).....	217
Tabla 33. Indicadores de Bondad de Ajuste del Modelo de Cinco Factores con Instrucciones de Respuesta que Inducía Honestidad al QI5F_tri, con un Diseño Entre-Grupos ( $N = 449$ ) .....	218
Tabla 34. Matriz de Cargas Factoriales Rotadas del Análisis Factorial Exploratorio del QI5F_tri Contestado Bajo Instrucciones de Faking, con un Diseño Entre-Grupos ( $N = 202$ ).....	219
Tabla 35. Indicadores de Bondad de Ajuste del Modelo de Cinco Factores con Instrucciones de Respuesta que Inducían Faking al QI5F_tri, con un Diseño Entre-Grupos ( $N = 202$ ) .....	220
Tabla 36. Matriz de Cargas Rotadas de Cinco Componentes Principales en la Condición Honesta para la Muestra Total ( $N = 939$ ).....	222
Tabla 37. Matriz de Cargas Rotadas de Cinco Componentes Principales en la Condición Faking para la Muestra Total ( $N = 692$ ).....	224
Tabla 38. Coeficientes de Ajuste del Modelo del Análisis de Componentes Principales .....	225
Tabla 39. Matriz de Cargas Rotadas de Cinco Componentes Principales en la Condición Honesta para la Muestra Intra-Grupo ( $N = 490$ ) .....	226



Tabla 40. Matriz de Cargas Rotadas de Cinco Componentes Principales en la Condición Faking para la Muestra Intra-Grupo ( $N = 490$ ) .....	227
Tabla 41. Coeficientes de Ajuste del Modelo del Análisis Factorial Exploratorio de la Muestra Intra-Grupo .....	228
Tabla 42. Coeficientes de Congruencia entre los Factores de Personalidad de la Condición Honesta ( $N = 939$ ).....	229
Tabla 43. Coeficientes de Congruencia entre los Factores de Personalidad de la Condición Faking ( $N = 692$ ).....	229
Tabla 44. Coeficientes de Congruencia entre los Factores de Personalidad de la Condiciones Honesta y Faking .....	230
Tabla 45. Coeficientes de Ajuste del Modelo del Análisis Factorial Confirmatorio para la Muestra Total.....	231
Tabla 46. Coeficientes de Ajuste del Modelo del Análisis Factorial Confirmatorio para la Muestra Intra-Grupo .....	232
Tabla 47. Características Demográficas de la Muestra ( $N = 939$ ).....	242
Tabla 48 A. Medias, Desviaciones Típicas y Correlaciones Observadas entre las Variables de la Muestra Total en la Condición Honesta .....	248
Tabla 49 A. Medias, Desviaciones Típicas y Correlaciones Observadas entre las Variables del Diseño Inter-Grupos en la Condición Honesta .....	252
Tabla 50 A. Medias, Desviaciones Típicas y Correlaciones Observadas entre las Variables del Diseño Intra-Grupo en la Condición Honesta .....	255
Tabla 51 A. Medias, Desviaciones Típicas y Correlaciones Observadas entre las Variables del Diseño Intra-grupo en la Condición Faking.....	258
Tabla 52. Fiabilidad y Valor de Restricción en el Rango ( $\alpha$ ) de los Big Five en Condición Honesta y Faking y Fiabilidad de las Medidas Criterio .....	261
Tabla 53. Correlaciones entre los Big Five en Condiciones Honesta y Faking con el Desempeño Global Empleando la Muestra del Diseño Intra-Grupo.....	262
Tabla 54. Correlaciones entre los Big Five en Condiciones Honesta y Faking con el Desempeño Académico Total Empleando la Muestra del Diseño Intra-Grupo.....	263





## ÍNDICE DE FIGURAS

Figura 1. Teoría de Snell et al. (1999). .....	105
Figura 2. Teoría de McFarland y Ryan (2000). .....	107
Figura 3. Teoría de Mueller-Hanson et al. (2006). .....	109
Figura 4. Teoría de Goffin y Boyd (2009). .....	111
Figura 5. Teoría de la Clasificación. ....	112
Figura 6. Teoría de Marcus (2009). .....	114
Figura 7. Teoría VIE (2011). .....	118





# INTRODUCCIÓN





La manipulación de las respuestas a los test no cognitivos por los evaluados es un tema de crítica relevancia para la Psicología del Trabajo y las Organizaciones que ha sido objeto de interés de la investigación durante las últimas décadas debido al creciente uso de medidas de personalidad en contextos organizacionales (Griffith y Peterson, 2006; Ziegler, MacCann y Roberts, 2012). Este comportamiento se ha denominado de diversas maneras, entre las más extendidas deseabilidad social o manejo de impresiones. No obstante, actualmente, el término más aceptado es el de *faking* (falseamiento o distorsión). Por ello, a lo largo de esta tesis doctoral, esta será la denominación que utilizaremos para referirnos a este fenómeno.

En el estudio del *faking* podemos definir claramente dos paradigmas desde los que se ha examinado este comportamiento. Por un lado, una perspectiva que ha tratado el *faking* como un fenómeno apenas existente o relevante y cuyos efectos son mínimos o inexistentes para las herramientas y procesos de evaluación (por ejemplo, Ellingson, Sackett y Connelly, 2007; Ellingson, Smith y Sackett, 2001; Hogan, Barret y Hogan, 2007; Hough, Eaton, Dunnette, Kamp y McCloy, 1990). Por otro lado, los estudios que han considerado el *faking* como un problema real en las herramientas de evaluación no cognitivas, particularmente, en los cuestionarios de personalidad (Birkeland, Manson, Kisamore, Brannick y Smith 2006; Donovan, Dwigth y Hurtz, 2003; Griffith, Chmielowoski y Yoshita, 2007; Salgado, 2016; Viswesvaran y Ones, 1999, entre otros). Actualmente, la evidencia empírica apoya el segundo planteamiento: los sujetos modifican sus puntuaciones en los tests a través del *faking* lo que, por consiguiente, afecta a todo el proceso evaluativo (Salgado, 2016). Asimismo, no se trata de un fenómeno aislado, se sabe que cualquier organización, con independencia del sector, que utilice cuestionarios de personalidad u otras medidas no cognitivas puede ver afectados sus procesos de selección y evaluación por el *faking*, al modificarse el ranking de candidatos y dar lugar a decisiones erróneas (Griffith y Converse, 2012).

Estos hallazgos han despertado el interés de los investigadores por desarrollar medidas que traten de controlar este problema. No obstante, la naturaleza intrínseca del *faking* ha hecho de este objetivo una tarea compleja, dando lugar a un conjunto de herramientas de evaluación con diferente grado de efectividad. Entre ellas, los cuestionarios de elección forzosa (*Forced-Choice Questionnaires*; de aquí en adelante EF) destacan como una medida capaz de reducir los efectos del *faking*. Se trata de

cuestionarios que obligan a los sujetos a seleccionar entre ítems de similar nivel de deseabilidad social, haciendo más difícil por ello que los evaluados respondan de modo socialmente deseable (Christiansen, Burns y Montgomery, 2005; Jackson, Wroblewski y Ashton, 2000).

En esta tesis se ha utilizado el modelo psicométrico del faking como el marco teórico para considerar los efectos del faking. Este modelo se ha puesto a prueba con los cuestionarios de personalidad tradicionales (*single stimulus*; de ahora en adelante SS), pero no con los cuestionarios de EF. El modelo propone que si se produce faking se observa un aumento de la media, una disminución de la desviación típica y, en consecuencia, se ve afectada la fiabilidad, la estructura factorial y la validez predictiva de los instrumentos. Los resultados con los cuestionarios SS han apoyado las predicciones del modelo (Salgado, 2016).

Esta tesis doctoral se presenta como una contribución empírica que pretende poner a prueba las predicciones del modelo psicométrico del faking en el caso de los cuestionarios de EF y estimar la efectividad de estas medidas contra el faking. Específicamente, los objetivos que se desean alcanzar son los siguientes: (a) estimar la robustez de los cuestionarios de EF contra los efectos del faking en la media y la desviación típica; (b) estimar los efectos del faking sobre la validez de constructo de los cuestionarios de personalidad de EF quasi-ipsativos y (c) estimar los efectos del faking sobre la validez de criterio de los cuestionarios de EF quasi-ipsativos.

Esta tesis se divide en dos grandes bloques para satisfacer los objetivos planteados. El primer bloque, el marco conceptual, tiene como fin delimitar el concepto de faking y realizar una revisión de las principales teorías sobre los antecedentes y consecuencias de este comportamiento y sobre las herramientas que se han desarrollado para su detección y control. Se compone de cinco capítulos teóricos que se detallan a continuación:

Titulado "*Faking: concepto y características*", el primer capítulo de esta tesis doctoral resume los aspectos más relevantes del concepto de faking. Se realiza una revisión de la evolución del concepto de faking y se detallan las características que definen este comportamiento. También se describen las principales taxonomías y se realiza un meta-análisis de la prevalencia del faking en los procesos de selección, siendo la primera contribución de esta tesis doctoral. Asimismo, se describen los principales métodos utilizados para operativizar el faking en la investigación, información que hemos

considerado esencial para poder interpretar los resultados empíricos descritos en esta tesis doctoral.

El segundo capítulo, *“Personalidad y faking”*, aborda las consecuencias que provoca el faking centrándose en las medidas de personalidad dado que son la herramienta, tras las entrevistas de selección, más utilizada y la que más sufre los efectos del faking. Primero se incluye una definición del concepto de personalidad y se realiza un resumen de las principales teorías para seguidamente describir con mayor detalle el modelo de los Cinco Grandes factores de personalidad en el que se centra esta tesis doctoral. Tras realizar una revisión de los principales hallazgos sobre la validez predictiva, a nivel ocupacional y educativo, de las medidas basadas en los Cinco Grandes, se describen los hallazgos más importantes de las consecuencias que provoca el faking.

En el tercer capítulo de esta tesis doctoral, titulado *“Modelos teóricos explicativos del faking”*, se exponen las principales variables que se han analizado como antecedentes de las conductas faking. Seguidamente se describen los principales modelos teóricos que han tratado de explicar los antecedentes del faking. Por último, se profundiza en las principales teorías que han tratado de explicar los efectos que produce el faking sobre las medidas de personalidad.

En el cuarto capítulo, *“Métodos de detección y control del faking”*, se realiza un repaso de las principales herramientas desarrolladas para detectar y disminuir la incidencia y efectos de este fenómeno. Se describe cada uno de los métodos y se detallan sus ventajas y limitaciones en la práctica profesional.

Para finalizar, el capítulo quinto, *“Cuestionarios de elección forzosa y el modelo psicométrico de los efectos del faking”*, se centra específicamente en esta herramienta por haber sido considerada una de las más adecuadas para la reducción de faking. Este capítulo incluye una descripción detallada de sus características, tipología, propiedades psicométricas y validez predictiva y de constructo. Para finalizar, se realiza un repaso de los principales estudios empíricos que han analizado la resistencia de los cuestionarios de EF al faking.

En el segundo bloque se incluye la parte empírica y, por tanto, las contribuciones únicas que se realizan en esta tesis doctoral. Se compone de tres estudios cuyo contenido se resume a continuación:

El primero de los estudios, titulado *“Meta-análisis de los efectos psicométricos del faking sobre los cuestionarios de personalidad de elección forzosa”* tuvo como objetivo realizar un meta-análisis de los estudios primarios que han analizado la

resistencia al faking de estas medidas de personalidad. Asimismo, se hizo un meta-análisis jerárquico para analizar variables moderadoras que puedan estar afectando a los resultados, en este caso, el tipo de diseño del estudio y el tipo de cuestionario de EF. Todos los meta-análisis llevados a cabo se realizaron siguiendo el método de meta-análisis experimental de diferencias con corrección de artefactos desarrollado por Hunter y Schmidt (1990, 2004; Schmidt y Hunter, 2015). Esta tesis es el único estudio empírico hasta el momento en el que se ha realizado un meta-análisis jerárquico de moderadores del faking y en el que se aplican correcciones por error de medida.

El segundo estudio incluido en esta tesis, que lleva por título *“Efectos del faking sobre la estructura factorial de los cuestionarios de personalidad de elección forzosa quasi-ipsativos: invarianza/equivalencia de la medida”* tuvo como objetivo examinar si el faking afecta a la estructura factorial de esta medida de personalidad de EF. Esta cuestión se examinó en una muestra general, en un diseño inter-grupos y en un diseño intra-grupo. Se realizaron análisis factoriales exploratorios (método de máxima verosimilitud), análisis de componentes principales y análisis factoriales confirmatorios para comprobar el ajuste de los datos al modelo de cinco componentes propuesto. Asimismo, se calcularon los coeficientes de congruencia de Burt-Tucker para obtener un índice más robusto de la similitud entre estructuras factoriales de la condición honesta y la condición faking.

El tercer estudio de esta tesis, *“Efectos del faking sobre la validez predictiva de los cuestionarios de personalidad de elección forzosa quasi-ipsativos”*, tuvo como principal objetivo determinar si las medidas de EF predicen el desempeño y los resultados académicos en condiciones de faking. Para ello se llevó a cabo un análisis de correlaciones para estimar la validez predictiva de estas medidas en condiciones de respuesta honesta y de respuesta inducida de faking. Además, en este estudio también se examinó el efecto moderador del tipo de medida (escalas de valoración vs. datos de resultados) y se ponen a prueba dos de las predicciones de la teoría psicométrica del faking, la reducción de la fiabilidad y de la validez predictiva (Salgado, 2016).

Como cierre, se presentan las conclusiones generales de los estudios empíricos realizados. Se resumen las aportaciones que se derivan de esta tesis doctoral y las principales implicaciones que tienen los resultados obtenidos para la teoría y la práctica profesional. Por último, se enumeran las limitaciones encontradas en los estudios realizados y se proponen algunas sugerencias para futuras investigaciones sobre el faking y los cuestionarios de EF.



Las referencias utilizadas para la redacción de la tesis y para los meta-análisis junto con los anexos figuran al final de esta memoria, lo que puede permitir a otros investigadores comprobar (re-analizar) nuestros resultados o servirse de ellos para futuros estudios.





# MARCO TEÓRICO

APROXIMACIONES TEÓRICAS, EFECTOS,  
MEDICIÓN Y CONTROL DEL FAKING



# CAPITULO 1

## FAKING: CONCEPTO Y CARACTERÍSTICAS





El faking es uno de los fenómenos más importantes a la vez que controvertidos de la psicología del trabajo y las organizaciones (de aquí en adelante PTO) y, particularmente, de la SP (de aquí en adelante SP). Existen diversas definiciones, planteamientos teóricos y consecuencias que se han asociado a este comportamiento a lo largo del tiempo, a pesar de que actualmente existe un amplio consenso en cuanto a los efectos negativos que provoca sobre las medidas no cognitivas, especialmente, sobre las medidas de personalidad.

Por esta razón, este primer capítulo presenta una revisión de la evolución del concepto de faking, cómo se ha medido y su prevalencia en los procesos de selección.

## 1. FAKING: CONCEPTO

La definición de faking es esencialmente compleja. Una revisión de la literatura científica sobre este fenómeno nos muestra las múltiples perspectivas desde las que se ha enfocado su estudio, bien como un rasgo de personalidad (Griffith, Malm, English, Yoshita y Gujar, 2006; Burns y Christiansen 2011), un sesgo en las respuestas (Paulhus, 1986, 2002), un estilo de respuesta (Ziegler, Toomela y Buehner, 2009) o inclusive desde una perspectiva puramente psicométrica (Zickar, Gibby y Robie, 2004). Además, han sido numerosos los términos con los que se le ha denominado o relacionado. Principalmente, se ha equiparado a conceptos como respuesta distorsionada (*response distortion*), respuesta socialmente deseable (*social desirable responding*), manejo de impresiones (*impression management*), deseabilidad social (*social desirability*), falsificación (*misrepresentation*) o engaño (*deception*). Pero también se han utilizado términos como sesgo de respuesta (*response bias*), conjuntos de respuesta (*response sets*), estilos de respuesta (*response styles*), mostrar virtudes improbables (*displaying unlikely virtues*) o autoensalzamiento (*self-enhancement*) para referirse a este mismo fenómeno (Griffith et al., 2006; Griffith, Chmielowski y Yoshita, 2007; Hough, et al., 1990; Mesmer-Magnus, Viswesvaran, Deshpande y Joseph, 2006; Ones, Viswesvaran y Reiss, 1996; Paulhus, 2002; Robie, Tuzinski y Bly, 2006; Ziegler, et al, 2012; Zickar y Gibby, 2006; entre otros).

Todos estos términos parten de la concepción del faking como la acción de modificar las puntuaciones en las pruebas evaluativas (Griffith y McDaniel, 2006; Robie et al., 2006). Aunque existe una tendencia a utilizarlos indistintamente, lo cierto es que existen ciertos matices que los diferencian entre sí. Por ello, consideramos de especial

relevancia comenzar este capítulo definiendo qué es el faking y diferenciándolo de otros conceptos a los que ha estado estrechamente ligado a lo largo de las investigaciones, particularmente, de los conceptos de deseabilidad social y manejo de impresiones.

Uno de los principales puntos de debate en el estudio del faking ha sido la propia naturaleza de este fenómeno. Aunque la perspectiva más extendida es que se trata de un comportamiento intencionado, existe una gran controversia sobre si es un proceso consciente o inconsciente. Los estudios más tempranos, en los años 30, que han planteado la posibilidad de que las respuestas a los test de personalidad se puedan distorsionar, únicamente han sugerido que esta conducta sólo se produciría cuando las personas tuviesen una motivación para hacerlo (Kelly, Miles y Terman, 1936; Vernon, 1934). No es hasta una década después cuando Meehl y Hathaway (1946) realizaron una importante diferenciación entre dos tipos de motivación, motivación consciente (faking o *lying*, según sus términos) y autoengaño inconsciente (*unconscious self-deception*). Los autores describieron el primero como un comportamiento relativo a personas que tienen disfunciones morales ya que voluntariamente falsean sus respuestas para alcanzar un objetivo personal. Por su parte, el autoengaño inconsciente es inherente a la persona y reflejo de un comportamiento natural en el que los evaluados ensalzan sus respuestas debido a que al hacerlo se sienten mejor consigo mismos. Meehl y Hathaway (1946) fueron, por tanto, los primeros en señalar que si las personas lo deseaban podían voluntariamente distorsionar sus respuestas. Sin embargo, fue Edwards (1957) el primero en usar el término *deseabilidad social* para referirse a la tendencia de las personas a mostrarse de un modo socialmente aceptable. Muchas investigaciones han equiparado la respuesta de faking a la respuesta socialmente deseable (p. ej. Ellingson, et al., 2001; Ones y Viswesvaran, 1999). Sin embargo, esta comparación no es lo suficientemente precisa en el contexto de la SP (Levashina y Campion, 2007).

El concepto de deseabilidad social ha sido extensamente estudiado y diversas investigaciones han puesto de manifiesto que se trata de un concepto multidimensional (Messick, 1960). En ese sentido, replicando la propuesta dual de la motivación de Meehl y Hathaway (1946), han sido varios los estudios factoriales que han analizado la deseabilidad social como un constructo compuesto por al menos dos factores. Wiggins (1964) fue uno de los primeros autores en explorar empíricamente la estructura de la deseabilidad social. Sus hallazgos dieron apoyo empírico a que este constructo tiene dos componentes, a los que llamó *factor alfa* y *factor gamma*. Ambos factores se



corresponden conceptualmente con los denominados posteriormente como autoengaño y manejo de impresiones, respectivamente (Paulhus, 1991, 2002).

Esta estructura dual ha sido replicada en estudios posteriores y con diversas denominaciones. Damarin y Messick (1965) distinguieron entre *sesgo autista*, entendido como la manifestación de una visión distorsionada de uno mismo y *sesgo propagandístico*, que los autores consideran como una tendencia ingenua a proyectar una imagen aceptable socialmente. Sackeim y Gur (1978), desde un enfoque puramente racional, diferenciaron la existencia de dos posibles procesos a la hora de responder a un cuestionario de personalidad, *autoengaño* y *engaño hacia otros*. En el primero, los sujetos reportan, convencidos de su veracidad, respuestas excesivamente positivas. El segundo se trata de una distorsión deliberada de la realidad. Ambas clasificaciones distinguen los componentes de la deseabilidad en base al nivel de consciencia de la conducta.

Sin embargo, no es hasta los estudios de Paulhus, (1984, 1986, 1991) cuando se reafirma la concepción de la deseabilidad social como un constructo compuesto por dos factores. Paulhus llevó a cabo tres estudios experimentales que consistieron en un análisis factorial exploratorio y confirmatorio y un estudio experimental del modelo bifactorial. Los resultados de estas investigaciones aportaron evidencia de que los dos factores que componen la deseabilidad social representan los constructos *autoengaño* y *manejo de impresiones*. Paulhus (1986) definió el autoengaño como “*la tendencia de las personas a describirse de un modo más positivo, pero sin ser conscientes de ello*” (pág. 144). Es decir, realmente creen que son como se describen, por lo tanto, es una conducta intrínseca al individuo y derivada de su autopercepción. El segundo componente, *manejo de impresiones* se refiere a “*la distorsión intencionada de la propia imagen para crear una impresión favorable*” (pág. 144). Por lo tanto, es un comportamiento consciente y temporal para alcanzar un fin concreto. Si lo aplicamos a un contexto de selección, cuando un candidato responde a las pruebas seleccionando la respuesta más deseable y no la más precisa sobre sí mismo está usando el manejo de impresiones. Por lo tanto, las respuestas podrían variar en función de las demandas sociales y la motivación (Paulhus, 1991).

Actualmente, la terminología de Paulhus (1991) es la más ampliamente aceptada y extendida, si bien existen otras conceptualizaciones del modelo de dos factores que se pueden enmarcar dentro de ésta. En la Tabla 1 se presenta un resumen de las más relevantes.

Tabla 1. Conceptualización de los Tipos de Respuesta Distorsionada

<i>Autores</i>	<i>Autoengaño</i>	<i>Manejo de Impresiones</i>
Damarin y Messick (1965)	Sesgo Autista	Sesgo Propagandístico
Sackeim y Gur (1978)	Creencias contradictorias; Autoengaño (inconsciente)	Engaño a otros (consciente)
Paulhus (1984)	Autoengaño	Manejo de impresiones
Paulhus (1986)	Intento inconsciente de mostrar una mejor imagen; Sesgo autista; Motivación inconsciente; Mecanismo defensivo inconsciente; Sinceridad	Falsedad; Sesgo Propagandístico; Insinceridad
Furnham (1986)	-	Faking; Simulación; Deseabilidad social; Engaño deliberado a otros
Paulhus (1988)	Ensalzamiento autoengañoso; Negación autoengañoso	Manejo de impresiones
Furnham (1990)	-	Distorsión motivacional o faking para alcanzar un perfil específico
Paulhus y Reid (1991)	Autodescripción honesta pero sesgada	Autodescripción favorable para otros
Paulhus (1991)	Autoengaño	Manejo de impresiones
Holden, Kroner, Fekken y Popham (1992)	-	Sensibilidad a los efectos de la audiencia Congruencia con los estereotipos
Seisdedos (1993)		Adaptación inteligente
Paulhus y John (1998)	Inconsciente defensa de la personalidad	Consciente defensa de la personalidad
Stark, Chernyshenko, Chan, Lee y Drasgow (2001)	-	Faking (distorsión intencionada de las respuestas a los ítems)
Paulhus (2002)	Alpha; Autoengaño	Gamma; Manejo de impresiones
Helmes y Holden (2003)	-	Sensibilidad a las señales sociales
Griffith, Chmielowski y Yoshita (2007)	Autopresentación exagerada	Faking; Distorsión intencionada

*Nota.* Tabla adaptada de “Seven nested questions about faking on personality tests”, de P. Tett, M. G. Anderson, C. Ho, T. S. Yang y A. Hanvongse, 2006. En R. L. Griffith y M. H. Peterson (Eds.), *A closer examination of Applicant faking behavior*, págs. 48-50, Greenwich, CT: IAP.

Tradicionalmente se ha identificado el faking con el concepto de manejo de impresiones, el segundo componente de la deseabilidad social (Levashina y Campion,

2007). Así, el faking se podría definir como la tendencia de los sujetos a modificar de modo voluntario sus respuestas a los instrumentos de evaluación (p. ej. test de personalidad, entrevistas, biodata, etc.). Las personas que comenten faking tratan de dar una mejor imagen de sí mismas que les permita obtener valoraciones más positivas. Concretamente, McFarland y Ryan (2000) consideran el faking como *“la modificación de las respuestas para mostrar una imagen más favorable”* (pág. 813) y Paulhus (2002) indica que se trata de dar una descripción *“excesivamente positiva”* (pág. 50). Por su parte, Ziegler et al. (2012) exponen una de las más recientes definiciones de faking como *“un conjunto de respuestas destinadas a proporcionar una imagen de uno mismo que ayude a lograr los objetivos personales. La falsificación se produce cuando este conjunto de respuestas se activa por demandas situacionales y características de la persona para producir diferencias sistemáticas en las puntuaciones de las pruebas que no se deben al atributo de interés”* (pág. 8).

Teniendo en cuenta estas definiciones, las conductas de faking están afectadas por variables sociales y situacionales, por lo que se puede entender que este fenómeno está mucho más próximo al concepto de manejo de impresiones exclusivamente y no al constructo global de deseabilidad social. No obstante, el constructo manejo de impresiones es más social y engloba todos aquellos comportamientos en los que las personas tratan de aparentar intencionadamente una imagen más deseable que su verdadera manera de ser, por ejemplo, en eventos sociales (Paulhus, 2012). Mientras que el uso del término faking se restringe exclusivamente al contexto de selección y evaluación y la distorsión de puntuaciones en medidas no cognitivas.

Si se toma en cuenta toda la literatura que ha profundizado en el concepto de faking, se pueden considerar cuatro aspectos esenciales que predominan a través de los diferentes estudios y que permiten conceptualizarlo (Ziegler, Maaß, Griffith y Gammon (2015):

- a) El faking es un comportamiento, no un rasgo de personalidad.
- b) Las personas a través del faking proyectan una imagen exagerada o distorsionada de ellas mismas.
- c) Para que se produzca faking los sujetos deben estar motivados y orientados a alcanzar un objetivo concreto.
- d) Los rasgos personales y las variables contextuales afectan a las conductas de faking.

## 2. TAXONOMÍAS DEL FAKING

Cuando se habla de faking necesariamente se tiene que pensar en las conductas que lo representan. De nuevo, en este punto existe una gran diversidad de comportamientos que por sus características podrían incluirse bajo esta etiqueta. Cuando incurren en faking, los sujetos manejan la información de múltiples maneras para obtener el resultado deseado. En el contexto de la SP pueden responder pensando en las características que tendría el candidato ideal para el puesto, exagerando su experiencia, omitiendo información que les perjudicaría en el proceso o, por supuesto, pueden ofrecer información inventada (Levashina y Campion, 2007). En ese sentido, han surgido diferentes taxonomías que intentan clasificar los comportamientos de faking.

Sackett, Burris y Callahan, (1989) describieron una de las primeras clasificaciones y de las más utilizadas hasta la actualidad. Según estos autores podemos diferenciar tres tipos de faking:

1. *Faking bueno o positivo*. Cuando los sujetos tienden a dar una imagen mejorada de ellos mismos al realizar las pruebas. Este comportamiento es específico del contexto de selección o promoción en los que se desea obtener un puesto de trabajo o mejorar las condiciones del actual.

2. *Faking malo o negativo*. Las personas potencian una imagen degradada de sí mismas. Este tipo de falseamiento tiene un mayor vínculo con los procesos de enjuiciamiento, cuando los sujetos pretenden disuadir su pena o responsabilidad apoyándose en dar la impresión de tener una (falsa) carencia en sus capacidades o un problema psicológico que reduzca o evite sus condenas (Salgado, 2005, 2016).

3. *Falseamiento de un rol específico* en el trabajo. Se refiere a responder escogiendo las alternativas que más se adecuen al rol que otros consideran que debe desempeñar.

El faking positivo es el que centraría el mayor interés de la SP ya que haría expresamente referencia al faking que tiene lugar en los procesos de selección cuando los candidatos tienden a dar una imagen más favorable de ellos mismos para mejorar sus resultados dando lugar a una distorsión en sus puntuaciones en el proceso de selección. Por consiguiente, este tipo de faking afectarían negativamente a los procesos de evaluación, no sólo porque los sujetos podrían estar falseando sus respuestas, sino

también porque los profesionales estarían valorando puntuaciones manipuladas y, por lo tanto, perjudicando todo el proceso de evaluación.

Furnham (1990) planteó que se pueden distinguir tres tipos de faking dependiendo del objetivo con el que se realice:

1. Personas que *sabotean las puntuaciones*.
2. Personas que *modifican las puntuaciones* en base a un perfil específico.
3. Personas que son *incapaces de describir sus propias características*.

De las tres, la segunda sería la más común en los procesos de SP.

Levin y Zickar (2002) sugirieron la existencia de diferentes tipos de faking según los procesos que los antecedían y las posibles consecuencias. Concretamente diferenciaron entre:

1. Presentación del comportamiento (*presentation behavior*). Los candidatos sólo exponen la información que más les favorecerá durante el proceso de selección. Se trataría de estructurar la información priorizando aquella que más beneficia durante la evaluación.
2. Falsificación del comportamiento (*falsification behavior*). En este caso se aumentan las posibilidades de contratación falsificando los comportamientos.
3. Comportamiento oportuno (*bullshit*), rompiendo la dicotomía de sólo considerar dos tipos de faking, mentir o decir la verdad (aunque sea guionizada), surge esta tercera categoría que engloba aquellos comportamientos que exponen la información por conveniencia, sin tener en cuenta la verdad o falsedad de esta.

En esa misma línea, Zickar et al. (2004) llevaron a cabo un estudio analizando si existían diferentes modos de incurrir en faking, encontrando clases de faking cualitativamente diferentes, lo que sugiere que las personas utilizan distintos procedimientos para cometer faking. Estos autores utilizaron la técnica MM-IRT (*Mixed-Model Item Response Theory*) para analizar los patrones de respuesta de los sujetos que cometen faking. Realizaron dos estudios experimentales en los que compararon las respuestas de personas motivadas a cometer faking con grupos que contestaron honestamente al cuestionario. Los resultados mostraron que, aún controlando las instrucciones de los grupos experimentales, existen diferentes patrones de respuesta. Esto

posiblemente se deba a las características personales de los participantes, por ejemplo, su capacidad de cometer faking, su autoconocimiento o la interpretación que hacen de las instrucciones. En cualquier caso, los datos permitieron distinguir de modo genérico tres tipos de patrones de respuesta basándose en la prevalencia de la respuesta:

1. *Respuesta regular*. Los sujetos escogen entre todo el abanico de posibilidades, sin inclinarse por ningún extremo en particular.
2. *Respuesta extrema*. Los candidatos se decantan siempre por las respuestas más positivas.
3. *Respuesta leve*. Es la opción intermedia entre las anteriores, las personas eligen respuestas positivas, pero no escogen las más extremas.

Por su parte, Levashina y Campion (2007) propusieron otra taxonomía de las conductas más representativas de faking. Tras identificar y analizar 125 conductas, clasificaron los comportamientos de faking en tres categorías: creación de la imagen, protección de la imagen y captación de la simpatía.

1. *Creación de la imagen*. Englobaría comportamientos tales como embellecer, adaptar, mejorar el ajuste, construir o inventar para obtener una buena imagen como candidato.
2. *Protección de la imagen*. Está relacionado con omitir, enmascarar y distanciarse para defender su imagen como candidatos adecuados para el puesto.
3. *Captación de la simpatía*. Esta categoría incluye comportamientos de conformidad de opinión y realce de la organización o evaluador.

Más recientemente, Kim (2011) realizó una revisión de la literatura sobre el faking y expuso que se pueden diferenciar cuatro estrategias para cometer faking en base al grado de información que manejan los candidatos:

1. *Reproducción*. Es la estrategia de faking en la que los candidatos manejan una mayor cantidad de información para manipular sus respuestas. Comportamientos que estarían incluidos en esta estrategia son obtener previamente una copia de la prueba, tener experiencia con procesos similares, copiar de un compañero u obtener información de otras personas que han pasado el mismo proceso. Esta estrategia no implica tener un éxito asegurado, pero permite a los candidatos saber qué deben hacer.



2. *Aplicación de reglas.* Se refiere a aquellas circunstancias en las que los candidatos tienen menos información del proceso y se apoyan en ciertas reglas para responder a las pruebas. En muchos casos se basan en la intuición, buscar la respuesta moderada o pensar cuál es la opción más deseable.

3. *Extrapolación.* En esta estrategia los solicitantes responden siempre basándose en un prototipo ideal. Los candidatos a partir de la concepción de ese candidato ideal extrapolan las respuestas en cada proceso de selección.

4. *Estrategias de Adivinación.* Se producen en contextos en los que los candidatos no tienen información acerca de cómo deben modificar las respuestas. En este caso, los sujetos sopesan cada una de las alternativas y van descartando las opciones.

No obstante, estas categorías no son necesariamente excluyentes entre sí y un mismo candidato puede poner en práctica una o varias de ellas según la situación lo requiera.

La existencia de diferentes modos de incurrir en faking revela que existen diferencias individuales y características contextuales que pueden estar afectando a la respuesta de faking (Paulhus, 1991) y, a su vez, esto puede producir diferentes consecuencias sobre los resultados del proceso de selección y sobre las medidas utilizadas. No obstante, antes de profundizar en el análisis de estas cuestiones se considera necesario aclarar cómo se ha operativizado el faking y realizar una estimación de su prevalencia en los procesos de SP para mostrar la incidencia real que tiene este fenómeno en los contextos de evaluación y SP.

### 3. OPERATIVIZACIÓN DEL FAKING

Si definir el fenómeno del faking ha supuesto una tarea controvertida y laboriosa, establecer cómo se debe medir este constructo también ha sido una cuestión objeto de debate. Las dificultades que se han encontrado a la hora de evaluar este constructo se deben, en gran medida, a su naturaleza intrínseca y al diverso conjunto de comportamientos que lo conforman (Ellingson, et al., 2007).

En ese sentido, se han aplicado diferentes métodos para cuantificar la intensidad y la frecuencia del faking: (a) las comparaciones entre grupos; (b) las diferencias en las medias y (c) el uso de escalas de deseabilidad social.

### 3.1 COMPARACIONES ENTRE GRUPOS

Se trata de una medida directa del faking basada en comparar las respuestas de muestras de sujetos en condiciones faking y condiciones honestas (Mesmer-Magnus y Viswesvaran, 2006; Smith y McDaniel, 2012). Esta estrategia se puede llevar a cabo combinando dos tipos de características: (a) el tipo de diseño utilizado (inter-grupos o intra-grupo) y; (b) el contexto en el que se desarrolle (experimental o real).

Respecto al tipo de diseño se pueden diferenciar entre:

1. *Intra-grupo*. En los diseños intra-grupo los mismos participantes contestan al cuestionario en condiciones honestas y bajo instrucciones que incitan a cometer faking. La literatura ha mostrado el importante poder estadístico de este tipo de diseño, ya que con una misma muestra se obtienen el doble de observaciones del funcionamiento del grupo. Asimismo, permite comparar las diferencias individuales que se pueden producir dentro de un mismo grupo ante diferentes condiciones. No obstante, esta metodología no está exenta de críticas. La más destacada es la sugerencia compartida por varios autores de que los resultados pueden estar afectados por la experiencia pretest del sujeto (recordemos que cubren dos veces la misma prueba) además de por las propias características personales de cada evaluado (Cook y Campbell, 1979; Mesmer-Magnus y Viswesvaran, 2006; Viswesvaran y Ones, 1999).

2. *Inter-grupos*. En este caso, el estudio se realiza con dos muestras independientes. Uno de los grupos contesta a la prueba de forma honesta y el otro grupo responde siguiendo instrucciones de falseamiento.

Una cuestión que se plantea sobre estos dos tipos de diseños es si se obtienen resultados igual de veraces independientemente de cuál de ellos se aplique. Viswesvaran y Ones (1999) estudiaron la incidencia de la falsificación en las medidas de personalidad utilizando ambos diseños y compararon los estimadores resultantes en cada uno. Concluyeron que los diseños intra-grupo son mucho más precisos explicando los efectos del faking. La principal crítica a los diseños inter-grupos es que se asume que el efecto del faking será igual para todos los sujetos aun cuando no tenemos datos de esa interacción con parte de la muestra ya que en los diseños inter-grupos sólo parte de la muestra responde en condiciones faking. Además, en estos diseños no existe necesariamente una equivalencia entre los grupos. Por lo tanto, se asume que los diseños



intra-grupo tendrán un mayor poder estadístico al permitir equiparar totalmente ambos grupos.

En cuanto al contexto podemos diferenciar dos tipos:

1. *Estudios correlacionales o en contextos reales.* Se trata de la comparación de las puntuaciones entre muestras de candidatos y de ocupantes de puestos de trabajo. Este paradigma asume que las personas que ya tienen el trabajo estarán menos motivadas a modificar sus respuestas, por lo tanto, las diferencias en las medias de las puntuaciones que se observen entre ellos y los solicitantes indicarán si se ha cometido faking (Mesmer-Magnus y Viswesvaran, 2012). Existe otra variante de este procedimiento y es comparar muestras de trabajadores con alta motivación y baja motivación a cometer faking. En cualquiera de los casos, este método se puede considerar un indicador del desempeño típico, es decir, la diferencias que se producen de modo natural en un contexto laboral normal. No obstante, dado el complejo planteamiento de este tipo de investigaciones en contextos reales, la mayoría de los estudios se realizan en contextos experimentales en los que se simulan procesos de selección.

2. *Procedimientos experimentales.* Los estudios que siguen este método se realizan en contextos controlados de laboratorio. A través de las instrucciones que se facilitan a los participantes, estos son asignados a la condición *fake good* (equivalente al papel de candidato) o a la condición honesta (equivalente al ocupante del puesto) antes de responder a las pruebas. En el falseamiento positivo (*fake good*) los participantes son instruidos para que respondan al test tratando de mostrar una imagen mejorada de sí mismos durante la evaluación. En la condición honesta se les indica que deben ser totalmente sinceros al responder a los test. En este caso se considera que los resultados reflejan los efectos del faking en contextos de desempeño máximo en los que claramente se pretende encontrar la mayor diferencia entre ambas condiciones experimentales. También existe otro tipo de instrucciones, las referidas a la condición de faking negativo o *fake bad* pero es poco común que en el ámbito de la SP se le requiera al candidato que responda mostrando una imagen negativa de sí mismo, reduciéndose a contextos muy específicos.

Salgado (2016) realizó una revisión meta-analítica sobre los efectos del faking en la que diferenció entre ambos procedimientos. Los resultados alcanzados sugieren que el

faking en un contexto real y el faking en un contexto experimental (con un diseño intra-grupo) produce los mismos efectos. Esto tiene una importante implicación para la investigación empírica ya que se puede asumir que los resultados encontrados en estudios experimentales tienen la misma veracidad que los encontrados en contextos reales. Por lo tanto, el uso de ambos procedimientos en la investigación del faking estaría más que respaldado.

### 3.2 DIFERENCIAS DE MEDIAS

Se centra específicamente en analizar las diferencias en las puntuaciones medias que obtienen los sujetos en las pruebas. Se trata de una estimación de la incidencia del faking mediante la aplicación de intervalos de confianza (Griffith et al., 2007; Peterson, Griffith, O'Connell y Isaacson, 2008). El objetivo es identificar a las personas cuyo cambio en las puntuaciones excede los límites que se pueden esperar teniendo en cuenta el error de medida. Parte de la consideración de que las puntuaciones falsas son aquellas que se encuentran fuera de un intervalo de confianza del 95% asociado a la media de ese factor. Para calcular el límite superior del intervalo se han utilizado típicamente dos técnicas: (a) el cálculo a partir de error estándar de medida de la prueba (SEM; por ejemplo, Griffith et al., 2007) y (b) el cálculo a partir del error estándar de la diferencia (SED; Peterson et al., 2008). Una vez obtenido este valor, se crea un intervalo de confianza basado en la puntuación del contexto de investigación, por ejemplo, se multiplica por 1.96 para obtener un intervalo de confianza del 95%. El resultado es el establecimiento de lo que esencialmente podría denominarse una puntuación de corte para el faking. Si la puntuación de un individuo excede el límite superior de este intervalo de confianza se considera un “*faker*”. De este modo, al conocer que personas han mentado, se puede estimar la incidencia del faking.

### 3.3. ESCALAS DE DESEABILIDAD SOCIAL

Esta metodología parte de la consideración de que la mejor manera de saber si las personas distorsionan sus respuestas es administrando, junto con las medidas de personalidad, una medida que evalúe las respuestas deseables. Se asume que las personas que obtienen puntuaciones altas en estas escalas están exagerando sus respuestas en toda la medida de personalidad. Una alternativa a este método son las escalas que incorporan ítems “*bogus*”, es decir, preguntas que incluyen información que no tiene relación con el

puesto de trabajo o tareas falsas que no se realizan en el puesto de trabajo. Si los sujetos eligen estas respuestas se considera que están distorsionando los resultados. A pesar de lo extendido de esta práctica, no son pocas las críticas a este paradigma. Principalmente se pone en duda la adecuación de los criterios que se aplican para discernir cuando se está cometiendo faking de cuando se está respondiendo honestamente a las preguntas (Burns y Christiansen, 2006; Goffin y Christiansen, 2003; Smith y Ellingson, 2002).

Este método se ha utilizado, además de como una medida de análisis de la prevalencia de faking, como un método de detección de este fenómeno. Por ello, será analizada en mayor profundidad en capítulos posteriores de esta tesis doctoral.

#### 4. PREVALENCIA DEL FAKING

La literatura ha respaldado la existencia del fenómeno del faking. Pero esta afirmación genera una segunda cuestión y es ¿en qué medida realmente los candidatos cometen faking cuando participan en un proceso de evaluación? (Rees y Metcalfe, 2003). Para responder a esta pregunta es necesario obtener un estimador preciso de la prevalencia del faking (Donovan, Dwight y Hurtz, 2003). Sin embargo, apenas existen estudios que hayan centrado su atención en esta cuestión debido, en gran medida, a la dificultad que supone evaluar la incidencia del faking al tratarse, como ya hemos indicado, de un comportamiento interno o cognitivo (Griffith y Converse, 2012).

Con el fin de aclarar esta cuestión, Griffith y Converse (2012) realizaron una revisión narrativa de los principales estudios que han tratado de estimar el porcentaje de personas que cometen faking. En su trabajo revisan sobre todo estudios basados en lo que ellos denominan estudios de *evidencia directa* del faking, en el que se comparan los resultados de personas motivadas y no motivadas a cometer faking. En el ámbito laboral se traduciría en comparar las respuestas de trabajadores con candidatos en procesos de selección. Esto autores se centran, especialmente, en las investigaciones que utilizan diseños intra-grupo, ya que, como hemos indicado, la literatura ha puesto de manifiesto que este diseño es más riguroso (Viswesvaran y Ones, 1999).

Los estudios incluidos en esta revisión se resumen en la Tabla 2. De izquierda a derecha aparecen el tipo de diseño y autores, el contexto en el que se ha desarrollado la investigación, el tamaño de la muestra, el constructo sobre el que se evalúa la prevalencia del faking y el porcentaje de faking encontrado en cada una de las investigaciones.

Tabla 2. Revisión de Estudios sobre la Prevalencia del Faking de Griffith y Converse (2012)

Estudio	Contexto	N	Constructo	Porcentaje
<i>Intra-grupo</i>				
Griffith, Chmielowski y Yoshita (2007)	Real	141	Conciencia	22 ~ 31
Peterson, Griffith, O'Connell y Isaacson (2008)	Real	123	Conciencia	15 ~ 26
Donovan, Dwight y Schneider (2008)	Experimental	198	Logro	27
			Potencial	19
Arthur, Glaze, Villado y Taylor (2009)	Experimental	296	Amigabilidad	36
			Conciencia	34
			E. Emocional	33
			Apertura	15
Peterson, Griffith y Converse (2009)	Experimental	370	Conciencia	19
<i>Items "Bogus"</i>				
Anderson, Warner y Spencer (1984)	Real	351	Desempeño	45
Pannone (1984)	Real	221	Desempeño	35
Donovan, Dwight y Schneider (2008)	Real	198	Logro (Conciencia)	21
<i>Autoinforme</i>				
McDaniel, Douglas y Snell (1997)	Real	192	Curriculum	42
Donovan, Dwight y Hurtz (2003)	Real	743 (4 muestras)	Desempeño	29
Robie, Brown y Beaty (2007)	Experimental	12	Conciencia	25
			Amigabilidad	
<i>Otros: revisiones</i>				
Hough (1998b)	Revisión	-	Conciencia	39
			Amigabilidad	49
Aamodt (2003)	Revisión	-	Curriculum	25
Levashina y Campion (2007)	Compilación Estudios Experimentales	825	Entrevista de selección	49
<i>% promedio</i>				30.5

*Nota.* Tabla adaptada de "The rules of evidence and the prevalence of applicant faking", de R. L. Griffith y P. D. Converse, 2012. En M. Ziegler, C. MacCann y R. D. Roberts (Eds.), *New perspectives of faking in personality assessment*, págs. 34-52, Nueva York, NY: Oxford University Press.

Tomando en conjunto los resultados de estos estudios se puede concluir que:

1. Los resultados en todos los estudios reflejan claramente que un considerable número de sujetos distorsionan sus respuestas cuando responden una prueba evaluativa no cognitiva, principalmente test de personalidad. Los estudios revisados demuestran que no sólo intentan cometer faking, sino que al menos un 20% de los sujetos llegan a modificar realmente sus respuestas (Donovan, Dwight y Schneider, 2008). En algunos casos, incluso, casi la mitad de los participantes en el estudio incurren en faking (49%) como, por ejemplo, en los trabajos de Hough (1998b) y Levashina y Campion (2007). No obstante, hablamos de que, en promedio un 30% de las personas comete faking.

2. Se constata que los candidatos tienden a cometer faking en mayor medida que las personas que ya desempeñan un trabajo. En los resultados se observa que las muestras de candidatos tienen valores más elevados en sus medias que las muestras de trabajadores.

Esta revisión de la literatura abarca muchos de los estudios realizados hasta el año 2012, pero no son los únicos trabajos que han tratado de cuantificar el faking. Por ejemplo, Rosse, Stecher, Miller y Levin (1998) encuentran que al menos un 18% de las personas ( $N = 197$ ) cometen faking en procesos de selección y que, de nuevo, los candidatos abusan en mayor medida de este comportamiento. Asimismo, en los últimos años han surgido más investigaciones sobre la prevalencia de este fenómeno. Es el caso del trabajo de Donovan, Dwight y Schneider publicado en el año 2014. Estos autores consideraron que la mayoría de los estudios previos se han realizado en contextos de laboratorio y que apenas unos pocos han tenido en cuenta criterios organizacionales reales para llevar a cabo la recopilación de datos. Por ello, realizaron un estudio sobre la prevalencia del faking en un contexto real de selección. A través de un diseño intra-grupo compararon las puntuaciones de una muestra de 162 individuos antes y después de ser contratados en una compañía farmacéutica. Para el estudio utilizan como instrumento el *Dispositional Goal Orientation* (VandeWalle, 1997) que evalúa la orientación al logro de metas. Los resultados de este estudio siguen la línea de sus predecesores. Concretamente, encontraron que al menos un 49.7% de los participantes son categorizados como fakers en al menos una de las tres dimensiones evaluadas. Los resultados también muestran que en torno al 9.3% de la muestra tiene comportamientos

de faking más extremos, es decir distorsionan sus respuestas en la mayor parte de las dimensiones de la prueba. Bahri (2013) obtuvo resultados similares en una muestra real de candidatos a una escuela de medicina en Malasia ( $N = 162$ ). Encontró que un 50.8% de los sujetos distorsiona sus respuestas en al menos un factor de personalidad.

Bott, O'Connell, Ramakrishnan y Doverspike (2007) compararon los resultados de personalidad entre una muestra de candidatos (5,629) y una de empleados (303) y también encontraron diferencias significativas entre ambos grupos,  $d = 1.05$  para amigabilidad y conciencia y  $d = 0.66$  para estabilidad emocional. En términos de contratación, estas diferencias implicarían que se pasaría de un 22,4% de sujetos que cumplen los requisitos de contratación a un 62%, siendo estos resultados una evidencia más de que los candidatos tienden a aumentar sus puntuaciones y, por tanto, de que el faking se produce.

Alternativamente, otros autores han llevado a cabo estudios de opinión sobre los comportamientos de faking. Estos estudios no aportan porcentajes o datos concretos sobre la incidencia del faking pero permiten obtener una visión de las actitudes que existen hacia este comportamiento en los contextos en los que se evalúa, en nuestro caso las organizaciones. En términos generales estos trabajos analizan los puntos de vista y percepciones del faking en colectivos que pueden verse afectados por este fenómeno bien sea por ser potenciales participantes en procesos de selección o por ser personas que desempeñan trabajos en puestos de gestión de recursos humanos.

Rees y Metcalfe (2003) realizaron un estudio para analizar las opiniones sobre el faking de potenciales candidatos o usuarios de test de personalidad en el que entrevistaron a 190 trabajadores de diferentes departamentos de recursos humanos del noroeste de Inglaterra. Los participantes cubrieron un cuestionario anónimo que exploraba sus opiniones sobre los comportamientos de faking. Los resultados mostraron que la mayoría de los entrevistados consideran que es fácil para los candidatos distorsionar sus respuestas en los test de personalidad ( $r = .38$ ;  $p < 0.01$ ). Además, creen que es muy difícil detectar quién está cometiendo faking, lo que alienta a cometer faking a pesar de que se considera este comportamiento como poco ético ( $r = .39$ ;  $p < 0.01$ ). Estos resultados revelan el punto de vista de una muestra de potenciales candidatos sobre el faking y evidencian la facilidad con la que creen que se pueden distorsionar los resultados de las pruebas de personalidad. Por su parte, English, Griffith, Graseck y Steelam (2005) encuestaron a una muestra de solicitantes de trabajo sobre sus percepciones del faking y reportaron que el



74% de los solicitantes están convencidos de que los candidatos llevan a cabo comportamientos de faking.

Además, de la evidencia y prevalencia del faking en medidas de personalidad, también se ha constatado sus efectos sobre otras herramientas de selección no cognitivas (entrevistas y curriculum vitae; CV) y en otros contextos evaluativos. Aunque estos últimos no sean el foco de atención de esta tesis doctoral, se considera pertinente citar sus aportaciones.

Levashina y Campion (2007) estudiaron los efectos del faking sobre la entrevista de selección y encontraron que en torno al 90% de los participantes tienen comportamientos de faking y que al menos un 25% de los sujetos distorsionan severamente sus respuestas.

Respecto a la predominancia del faking en los CV, una revisión de la literatura muestra que, al menos, un 20% de los CV contiene modificaciones en los estudios cursados (Griffith y Converse, 2012). Igualmente, Aamodt (2003), tras examinar los datos proporcionados por 15 empresas y portales de empleo, concluyó que, en promedio, un 25% de los candidatos distorsiona el CV y Callahan (2004) encuentra que un 41% de las personas admite haber falsificado su CV. En conjunto, estas cifras evidencian que el faking no sólo se limita a los test de personalidad y medidas relacionadas, sino que se extiende en igual medida a otras herramientas no cognitivas popularmente utilizadas en los procesos de SP.

Si dirigimos nuestra atención hacia otros entornos de evaluación, podemos encontrar resultados similares sobre la prevalencia de faking. Así, Hall y Hall (2011) estudiaron la prevalencia del faking en contextos de evaluación médica y psicológica. Estimaron que, tanto si se trata de casos de lesiones físicas como de simulación de enfermedades mentales o evaluaciones criminales, alrededor del 30% de las personas distorsionaba sus respuestas, tal y como ocurre en contextos de selección.

En definitiva, en todos los estudios citados se han obtenido conclusiones similares a través de diferentes muestras, contextos y pruebas. Estos resultados permiten concluir que el faking ocurre y que, al menos, un tercio de las personas distorsionan sus respuestas. Sin embargo, todos los estudios que se han citado son investigaciones primarias que aportan datos de casos particulares de muestras reducidas o revisiones que incluyen datos de diversa naturaleza que pueden estar produciendo un sesgo en la interpretación de los datos. Por ejemplo, comparar muestras en contextos reales y experimentales o no diferenciar el tipo de medida sobre la que se está evaluando la incidencia del faking. Estas

limitaciones ponen de manifiesto la necesidad de llevar a cabo una revisión más exhaustiva de los estudios sobre la prevalencia del faking.

Con el objetivo de superar las limitaciones descritas, consideramos necesario terminar este capítulo realizando la primera contribución empírica de esta tesis doctoral con la realización de un meta-análisis para obtener conclusiones más robustas sobre la prevalencia real del faking en los procesos de evaluación y selección.

Por tanto, se llevó a cabo un meta-análisis *bare-bones* en el que, además de realizar un análisis general, se estudió el papel moderador de dos variables criterio sobre la prevalencia del faking (a) el tipo de diseño experimental y (b) el tipo de medida utilizada. Todos los estudios incluidos en este meta-análisis aportaban el valor del tamaño de la muestra y el porcentaje estimado de prevalencia. En los casos en los que se presentaban varios valores de prevalencia se procedió calculando un valor promedio para el constructo medido. Los datos obtenidos se resumen en la Tabla 3.

Los resultados obtenidos muestran que, independientemente de la variable criterio, un 37% de los sujetos cometen faking en los procesos de evaluación. No obstante, si no tenemos en cuenta las muestras más grandes, en este caso el estudio de Bott et al. (2007) cuya muestra representa más de la mitad de los datos recogidos y, por lo tanto, podría estar condicionando los resultados obtenidos, el valor de prevalencia se reduce hasta el 34%. Este valor es ligeramente superior al obtenido directamente de los estudios primarios, aunque en el caso de las medidas de personalidad y los estudios experimentales la prevalencia promedio no alcanza el 30%.

Tabla 3. Resultados del Meta-análisis de la Prevalencia del Faking

	Muestra total			Sin muestras grandes		
	<i>N</i>	<i>K</i>	% <sub>ow</sub>	<i>N</i>	<i>K</i>	% <sub>ow</sub>
Total	9820	16	37	4191	15	34
Diseño						
<i>Experimental</i>	1701	5	28	1701	5	28
<i>Real</i>	8119	11	38	2490	10	33
Medida						
<i>Personalidad</i>	7488	11	37	1859	10	27
<i>Desempeño</i>	2332	5	40	2332	5	40

*Nota.* *N* = tamaño muestral acumulado; *K* = número de muestras independientes; %<sub>ow</sub> = porcentaje de prevalencia observado ponderado por el tamaño de la muestra.



Si comparamos los resultados según el tipo de diseño, nos encontramos con un porcentaje mayor de prevalencia del faking en los procesos de selección reales, que oscila entre un 33% y un 38%, mientras que en los estudios experimentales no supera el 28%. Por su parte, centrándonos en el tipo de medida, los datos muestran, en conjunto, que entre un 27% y un 37% de las personas llegan a modificar sus puntuaciones en los test de personalidad y que este valor se agrava en medidas de desempeño hasta alcanzar el 40% de prevalencia. Aun así, hemos de indicar que los datos pueden estar condicionados por el tamaño de las muestras, ya que a medida que aumenta este valor también lo hace el porcentaje de prevalencia. En todo caso, estos resultados demuestran que al menos un tercio de los sujetos cometen faking en los procesos de SP independientemente del tipo de diseño o medida utilizada.

A lo largo de este capítulo hemos puesto de manifiesto la singularidad y complejidad de este fenómeno, así como lo extendido que se encuentra en la práctica profesional. Los datos revelan que el faking no se trata sólo de un controvertido constructo teórico, sino que los datos empíricos revelan que se produce en la práctica profesional de la SP y que, particularmente, afectan a los procesos de SP. En ese sentido, es de gran importancia el avance en el conocimiento de este fenómeno, especialmente en la estimación de su impacto real en las medidas y los procesos de evaluación, pero también en el desarrollo de estrategias que permitan contenerlo. Esta tesis pretende ser una contribución a estos objetivos.

De todo lo expuesto se pueden destacar las siguientes conclusiones:

1. El faking se refiere a distorsión voluntaria de las respuestas a las pruebas evaluativas con el fin de mostrar una imagen favorable que permita obtener los objetivos deseados. Por lo tanto, la principal característica que define al faking es que se trata de un comportamiento intencionado, es decir, voluntario.
2. Cuando se define el faking se debe diferenciar del concepto de deseabilidad social. Este último, consiste en un constructo más amplio que abarca todos los comportamientos de adaptación (consciente e inconsciente) a lo socialmente aceptable. Uno de los factores que lo componen es el manejo de impresiones, en donde se ubicaría el faking.
3. Han surgido varias taxonomías del faking. La más extendida y que utilizaremos a lo largo de esta tesis doctoral, diferencia dos tipos de faking

principalmente: faking positivo o ensalzamiento de la imagen y faking negativo o degradación de la imagen. El primero se produce en contextos de selección, el segundo se asocia principalmente a procesos judiciales.

4. Como primera contribución de esta tesis doctoral se realizó un meta-análisis para estimar la prevalencia del faking en función de varias variables criterio. Los resultados mostraron que, independientemente del tipo de medida utilizada y del diseño del estudio, en promedio, un 34 % de los sujetos cometerían faking.

5. Es necesario estimar qué efectos provoca este fenómeno sobre las medidas no cognitivas. Si conocemos cómo afecta el faking a las puntuaciones podremos desarrollar medidas específicas que lo corrijan.



## CAPITULO 2

# PERSONALIDAD Y FAKING





Como hemos descrito en el capítulo previo de esta tesis doctoral, el faking es un fenómeno asociado a las medidas de evaluación no cognitivas. En esta categoría podemos englobar un amplio conjunto de herramientas que se han venido utilizando con diferente grado de frecuencia en los procesos de selección y evaluación de personal. Entre ellas, podemos destacar, por ejemplo, los biodata, los assessment centers, las entrevistas o los cuestionarios de personalidad.

Es precisamente esta última, la personalidad, la que ha acaparado la mayor parte de la atención en lo que al estudio de los efectos del faking se refiere. Los cuestionarios de personalidad son una de las herramientas más extensamente utilizada en los procesos de SP (Rothstein y Goffin, 2006). Además, el modelo de personalidad de los Cinco Grandes, en el que se basan las pruebas más actuales, se sustenta en un marco teórico más robusto que el de cualquier otra medida, por lo que no resulta sorprendente que la investigación se haya centrado casi exclusivamente en esta medida para conocer los efectos reales que provoca el faking (Salgado, 2016).

Debido a la importancia de la personalidad en el estudio del faking, en este capítulo realizaremos un repaso en la evolución histórica del concepto de personalidad normal y su evaluación, analizaremos su validez como medida predictiva del desempeño y resumiremos la literatura científica que ha aportado evidencia empírica de los efectos que tiene el faking sobre las medidas de personalidad en los procesos de selección y evaluación.

## 1. PERSONALIDAD: MODELO DE LOS CINCO GRANDES

La personalidad normal es uno de los constructos más importantes y sobre el que más literatura científica se ha publicado en el ámbito de la psicología. Costa y McCrae (1989; Costa, McCrae y Kay, 1995) la definieron como *“los estilos relativamente estables de pensar, sentir y actuar que caracterizan a un individuo”*. Además, indicaron que estos rasgos *“son estables y se reflejan en muchos aspectos de la vida cotidiana de una persona incluyendo preferencias vocacionales, opciones de carrera y desempeño laboral”* (pág. 124). Por lo tanto, cuando hablamos de personalidad nos referimos a un conjunto de características relativas al individuo, estables en el tiempo y de naturaleza psicológica.

Diferentes modelos y teorías se han desarrollado para tratar de explicar este constructo desde una perspectiva psicológica. Entre ellos, podemos destacar el Modelo de 16 Factores de Cattell (Cattell y Eber, 1964) o el Modelo de 3 Factores de Eysenck (Eysenck y Eysenck, 1975). Sin embargo, desde los años 90, el modelo más dominante en la psicología de la personalidad es el Modelo de los Cinco Grandes o Big Five (Costa y McCrae, 1992; Hogan, 1991) el cuál define la personalidad en base a cinco grandes dimensiones.

Los comienzos de este modelo se remontan a los estudios iniciales de McDougall (1932) y de Allport y Odbert (1936) sobre la composición del constructo de la personalidad. McDougall fue el primero en proponer que la personalidad se podría definir en términos de cinco factores: templanza, carácter, intelecto, temperamento y disposición. Por su parte, Allport y Odbert (1936) elaboraron un conjunto de listados con términos lingüísticos que definían los rasgos estables de la personalidad. Si bien, no fue hasta los años 40 cuando Cattell (1943, 1946, 1947, 1948) desarrollaría los primeros análisis factoriales de la personalidad normal. Cattell, apoyándose en las listas de Allport y Odbert (1936), desarrolló varias escalas bipolares para la medición de la personalidad que, tras los pertinentes análisis factoriales, dieron lugar a un conjunto de 16 factores de primer orden y varios de segundo orden. Aunque se trataba de un modelo de una elevada complejidad, las escalas de Cattell fueron el punto de partida de relevantes investigaciones posteriores.

Fiske (1949), primero, y Tupes y Christal (1961), después, volvieron a analizar los datos aportados por Cattell, revelando una estructura factorial más clara que la obtenida por este autor. Fiske (1949) analizó las respuestas dadas por 125 participantes a 21 escalas bipolares de Cattell y halló cinco factores diferenciados, convirtiéndose en el primero en obtener claramente una estructura factorial de cinco componentes. Sin embargo, la investigación realizada por Tupes y Christal (1961) es la que se ha llevado el mérito de ser considerada como la primera que identificó los cinco factores de personalidad tal y como son definidos en la actualidad.

Esta estructura factorial ha sido replicada posteriormente por un extenso conjunto de investigadores, por ejemplo, Norman (1963a), Smith (1967) o Goldberg (1981). Los estudios volvieron a ratificar que la personalidad se conforma por cinco grandes factores. No obstante, es a partir de los años 90 cuando se demuestra la robustez de este modelo. Los estudios más recientes han mostrado que se trata del modelo más sólido para explicar la personalidad normal (Goldberg, 1981) y que, además, es generalizable en muestras de

diversos países y culturas a través de diferentes instrumentos de evaluación (p. ej. Borkenau y Ostendorf, 1990; Caprara y Perugini, 1994; McCrae y Costa, 1985, 1987, Salgado, 1997).

Como hemos indicado, el Modelo de los Cinco Grandes o Big Five, afirma que la personalidad abarca cinco grandes dimensiones a los que actualmente se les denomina estabilidad emocional, extraversión, apertura a la experiencia, amigabilidad y conciencia (Costa y McCrae, 1989).

1. *Estabilidad emocional*. En muchas ocasiones también es denominado con su polo opuesto *neuroticismo* o *inestabilidad emocional*. Este factor hace referencia al ajuste emocional de las personas. Los sujetos que puntúan alto en ese factor son personas que se caracterizan por el control emocional, la tranquilidad y la predominancia de sentimientos positivos, tales como alegría, serenidad o buen humor. Por el contrario, las personas inestables emocionalmente son personas con falta de control sobre sus emociones, tienden a ser inseguras, muy irritables y nerviosas y tienden a experimentar emociones negativas como, por ejemplo, odio, enfado, tristeza (Costa y McCrae, 1992; Salgado, 1998).

2. *Extraversión*. Refleja el carácter social de las personas, es decir, la relación de los sujetos con su entorno. Aquéllos altos en extraversión son propensos a las relaciones sociales, a la necesidad de tener la compañía de otros y al atrevimiento en situaciones sociales. También se caracterizan por ser personas asertivas y habladoras. Su polo opuesto, por tanto, abarcaría aquellas personas que prefieren la soledad al contacto social, son reservados, introspectivos y poco dependientes de otros (Costa y McCrae, 1992; Salgado, 1998). Para Hogan (1986) este factor se compone de dos dimensiones, ambición y sociabilidad. El primero engloba características como iniciativa, sugerencia o ímpetu. Sociabilidad, por su parte, hace referencia a personas exhibicionistas y expresivas.

3. *Apertura a la experiencia*. La imaginación activa, la sensibilidad estética, el gusto por la variedad o la curiosidad intelectual son elementos que se englobarían en este factor. Así, las personas abiertas a la experiencia se definen como imaginativas, originales o curiosas. Por el contrario, quien puntúa bajo en este factor se caracteriza por ser conservador, convencional, rutinario en sus acciones, falto de imaginación y con preferencia a lo sencillo y lo obvio (Costa y McCrae, 1992; Salgado, 1998).

4. *Amigabilidad*. La persona amigable se caracteriza por su altruismo, tiene una clara necesidad de ayudar a otros y considera que la ayuda le será devuelta. Las personas altas en este factor se desviven por ayudar a los demás, son dóciles, honestos, cooperadores y amables. En su polo opuesto, la hostilidad, el egocentrismo y el escepticismo hacia los demás son las características dominantes. Estas cualidades, que por sí mismas son interpretadas como negativas, pueden ser beneficiosas en ciertos contextos y una ventaja práctica ya que el pensamiento crítico y escéptico contribuye a la elaboración de valoraciones juiciosas más objetivas (Salgado, 1998). Se trata por tanto de personas más competitivas que cooperativas.

5. *Conciencia*. Por último, el factor conciencia, también denominado *responsabilidad* o *deseo de logro*, hace referencia la capacidad de auto-control. Agruparía características como la eficacia, el orden y la organización, el sentido del deber, la responsabilidad o la perseverancia. Este factor está asociado con el logro académico y ocupacional (Salgado, 1998). Las personas altas en este factor, por tanto, son concienzudas, rigurosas y se desviven por cumplir con sus obligaciones. Los que obtienen puntuaciones bajas en este factor son personas más perezosas en el cumplimiento de sus obligaciones, siendo más laxos e informales en sus comportamientos (Costa y McCrae, 1992; Salgado, 1998).

## 2. VALIDEZ PREDICTIVA DEL MODELO DE LOS CINCO GRANDES

La validez de las medidas de personalidad para predecir diversos criterios organizacionales ha centrado una importante parte de la literatura científica de la SP desde comienzos del siglo XX (Barrick, Mount y Judge, 2001; Rothstein y Goffin, 2006). Los primeros estudios sobre esta cuestión plantearon que la personalidad no resultaba ser un adecuado predictor del desempeño. La revisión narrativa de Guion y Gottier (1965) fue uno de los trabajos que mejor evidenció esta visión pesimista sobre las medidas de personalidad. Estos autores hallaron que sólo alrededor de un tercio de los estudios que incluyeron en su revisión evaluaban la validez predictiva de estas medidas, por lo que concluyeron que no se podía mostrar la validez de las medidas de personalidad para la SP.



Sin embargo, con el comienzo de la década de los 90 la perspectiva cambia, la consolidación del Modelo de los Cinco Grandes da lugar a numerosas investigaciones y meta-análisis que evidencian la validez de las medidas de personalidad como predictor de diversos criterios ocupacionales y académicos.

Debido a lo extenso de la literatura científica que se ha centrado en investigar esta cuestión, en este apartado resumiremos brevemente los trabajos empíricos más relevantes y las principales conclusiones obtenidas en relación con la validez predictiva de los Cinco Grandes.

## 2.1 MODELO DE LOS CINCO GRANDES FACTORES Y DESEMPEÑO OCUPACIONAL

Las medidas de personalidad se han utilizado desde hace décadas como una herramienta preferente para la evaluación de los candidatos en los procesos de selección. Por ejemplo, Heller (2005) encontró que al menos un 30% de las compañías americanas utilizaba test de personalidad para evaluar a los solicitantes de trabajo. Sin embargo, hasta hace poco ha existido cierta controversia en relación a si verdaderamente son una medida que permite predecir el desempeño de los candidatos (Rothstein y Goffin, 2006).

El meta-análisis de Hough et al. (1990) fue de los primeros trabajos en publicar resultados robustos sobre esta cuestión. Estos autores analizaron la validez predictiva de varios constructos de personalidad sobre seis criterios relacionados con el trabajo: resultados educativos (desempeño), éxito en la formación, implicación laboral, eficacia laboral, abuso de sustancias y los comportamientos delictivos. Los resultados mostraron que cada uno de los factores predecía diferentes aspectos. Los factores *Adjustement*, *Intellectance* (apertura a la experiencia) y *Dependability* (conciencia) correlacionaron con casi la totalidad de las variables analizadas, con valores superiores a  $r = .13$ . El factor *surgency* (extraversión) se relacionó con el desempeño en el ámbito educacional ( $r = .15$ ) y los comportamientos delictivos ( $r = -.29$ ). Por último, el factor *Agreeableness* (o amigabilidad) predeciría el éxito en la formación ( $r = .10$ ) y los comportamientos delictivos ( $r = -.31$ ). No obstante, se trata de valores no corregidos ni por falta de fiabilidad en las medidas ni por restricción en el rasgo por lo que los resultados podrían resultar aún más contundentes con los ajustes pertinentes. Si bien, estos resultados ya contribuyen a mostrar la validez predictiva de las medidas de personalidad.

Se puede notar que estas medidas no están basadas, aunque sí podrían considerarse equivalentes, en el modelo de los Big Five sino en la estructura de seis

factores propuesta por Hogan (1983). Fue un año más tarde cuando Barrick y Mount (1991) publican uno de los estudios de mayor relevancia sobre la validez predictiva de las medidas de personalidad basadas en los Big Five.

Barrick y Mount (1991) llevaron a cabo el meta-análisis más importante de la década de los 90 sobre las medidas de personalidad. Estos autores examinaron la relación entre los Big Five y el desempeño ocupacional en cinco categorías profesionales a través de tres variables criterio. Los resultados hallados mostraron que conciencia es el único predictor válido para todos los puestos y ocupaciones analizados en el meta-análisis, obteniéndose una  $\rho$  promedio de .22. Además, el rango de valores de  $\rho$  para este factor osciló entre .20 y .23 lo que refleja una importante consistencia de la validez predictiva de este factor en las diferentes categorías ocupacionales. Respecto a los restantes factores de personalidad, su validez varía en función de los puestos analizados. De este modo, extraversión obtuvo una correlación significativa con aquellos puestos orientados a la atención a terceros, en este caso, puestos de gerencia o ejecutivos ( $\rho = .18$ ) y ventas ( $\rho = .15$ ). Amigabilidad predice el desempeño también en puestos orientados a la gerencia y policías, ambos con un valor  $\rho = .10$  y, por su parte, el factor estabilidad emocional se relaciona significativamente con la categoría de policías ( $\rho = .10$ ). El meta-análisis mostró que apertura a la experiencia no resulta ser un predictor válido para ninguna categoría profesional analizada. Con relación a las variables criterio examinadas en este meta-análisis (eficacia en el trabajo, éxito formativo y datos del personal), los resultados siguen la misma línea de los descritos para las categorías profesionales. Conciencia vuelve a ser el factor que generaliza su validez a través de los tres criterios analizados,  $\rho$  promedio de .22. Respecto a los restantes factores de personalidad, destaca que se encontró que todos los factores son predictores válidos del éxito en la formación y que amigabilidad lo es también de la variable criterio datos de personal, que engloba características como el salario o la antigüedad en el puesto. No obstante, como los propios autores recalcan “*el resultado más importante es sobre la dimensión conciencia*” (Barrick y Mount, 1991, pág. 17) al mostrarse que su validez se generaliza a través de los diferentes puestos y criterios analizados.

Una década después, Barrick et al., (2001), llevaron a cabo un nuevo meta-análisis, pero en este caso se trató de una meta-análisis de segundo orden con el objetivo de resumir cuantitativamente los resultados de 15 meta-análisis realizados hasta esa fecha sobre la relación entre el modelo de los Cinco Grandes Factores de personalidad y el desempeño laboral. En esta investigación, analizaron un amplio conjunto de variables

criterio relativas al desempeño y todos los estimadores fueron corregidos por error de medida en el predictor y el criterio y por restricción en el rango para, como los propios autores indican, “*hallar la verdadera relación teórica entre los Big Five y el desempeño*” (pág. 13). Los resultados obtenidos fueron consistentes con los hallados por Barrick y Mount (1991), esto es, conciencia demostró ser el único factor que generaliza su validez para todos los puestos y variables criterio analizadas. Asimismo, en todos los casos este factor obtuvo los valores  $p$  más altos, entre .20 y .30. Los demás factores sólo generalizaron su validez para algunos criterios en algunas categorías ocupaciones. No obstante, los autores destacan el factor de estabilidad emocional debido a que es el único de los restantes que muestra correlaciones que difieren de cero en todos los criterios analizados, lo que demuestra su valía como predictor del desempeño.

Los resultados descritos fortalecen la relevancia del modelo de los Big Five para la predicción de diversos criterios organizacionales y consolidan al factor conciencia como el mejor predictor del desempeño. Sin embargo, como señaló Salgado (1997), los meta-análisis realizados hasta ese momento estaban basados sólo en muestras americanas y canadienses, por lo que existía la necesidad de comprobar si las conclusiones obtenidas no se verían afectadas por las diferencias culturales.

Para dar respuesta a esta cuestión, Salgado (1997) realizó un meta-análisis en el que sólo incluyó muestras europeas que examinaban la relación entre los Big Five y tres criterios organizacionales: desempeño en el trabajo, éxito en la formación y datos personales. Los resultados mostraron que los factores de conciencia y estabilidad emocional son los mejores predictores del desempeño independientemente del puesto y los criterios analizados. Los tamaños del efecto, corregidos por error de medida y restricción en el rasgo, fueron en promedio  $p = .25$  para conciencia y  $p = .19$  para estabilidad emocional. Destaca también que se encontró que apertura a la experiencia y amigabilidad generalizaban su validez sobre el criterio éxito en la formación ( $p = .26$  y  $p = .31$  respectivamente). Los resultados descritos demuestran que la validez predictiva de los Big Five es generalizable a diferentes países y culturas.

Las investigaciones descritas no son las únicas que se han llevado a cabo sobre la relación entre los Big Five y el desempeño. Se han publicado nuevos trabajos meta-analíticos que han contribuido a la consolidación de estos resultados, demostrando que las medidas de personalidad son una importante herramienta para la PTO, en general, y la SP, en particular, y destacando en especial la utilidad de los factores conciencia seguido de estabilidad emocional como predictores de importantes criterios organizacionales (p.

ej., Hurtz y Donovan, 2000; Salgado, 2003; Salgado y Táuriz; 2014; Salgado, Anderson y Táuriz, 2015).

## 2.2 MODELO DE LOS CINCO GRANDES FACTORES Y DESEMPEÑO ACADÉMICO

El contexto académico ha sido otro de los ámbitos en los que más se ha estudiado la validez predictiva de los Big Five. Este interés se podría justificar en los resultados de los meta-análisis de Roth, Bevier, Switzer y Schippmann (1996) y Kuncel, Hezlett y Ones (2004) sobre la relación entre desempeño académico y ocupacional. Roth, BeVier, Switzer y Schippmann (1996) mostraron que el desempeño académico predecía el desempeño ocupacional al encontrar una correlación corregida de  $\rho = .35$  entre ambas variables. Asimismo, Kuncel et al. (2004) concluyeron en un meta-análisis más reciente que los factores que determinan el desempeño académico son los mismos que describen el rendimiento en el contexto laboral. Estos hallazgos demuestran la importante relación existente entre ambas variables. Por lo tanto, cuanto más conocimiento tengamos sobre los determinantes del desempeño académico más información tendremos sobre el futuro desempeño laboral de los sujetos. En ese sentido, una parte de las investigaciones en el campo educativo han tratado de analizar si el modelo de los Big Five es capaz de predecir diferentes criterios de éxito académico. Los hallazgos a este respecto se aproximan a los obtenidos en el ámbito laboral.

El meta-análisis de Poropat (2009) es considerado uno de los estudios más importantes que han analizado la relación entre los Big Five y el desempeño académico (entendido como las calificaciones obtenidas por los estudiantes). Poropat (2009) encontró que el factor conciencia fue el mejor predictor de desempeño académico con una  $\rho = .22$ , seguido del factor apertura a la experiencia que obtuvo una  $\rho$  de .12. El autor aclara que, aunque estos valores se puedan considerar bajos, se aproximan a los alcanzados por medidas de inteligencia (p. ej. la prueba SAT o *Scholastic Aptitude Test*) que han sido empíricamente avaladas como predictores del desempeño académico.

Por su parte, Salgado y Táuriz (2014) publicaron uno de los meta-análisis más recientes sobre esta cuestión. En este caso, además de analizar la relación entre personalidad y desempeño académico, también examinaron si el tipo de medida de personalidad utilizada afectaba la relación entre ambas variables. Estos autores llegaron a la misma conclusión que Poropat (2009), conciencia es el factor que mejor predice el

desempeño académico en todos los casos. No obstante, matizan que su validez es superior si se utilizan medidas de personalidad de EF quasi-ipsativas ( $\rho = .21$ ).

### 2.3 MODELO DE LOS CINCO GRANDES FACTORES Y VALIDEZ AÑADIDA

La evidencia empírica ha demostrado que las medidas de personalidad basadas en los Big Five son un adecuado predictor del desempeño tanto académico como ocupacional. Sin embargo, en la práctica real, estos instrumentos suelen administrarse junto a otras medidas como las entrevistas o los test de inteligencia. Se presupone que la utilización conjunta de medidas de diversa naturaleza proporcionará una mayor cantidad de información sobre los criterios de interés. Nos referimos al concepto de validez añadida de las medidas, es decir, a la magnitud de la contribución de un predictor tras conocer la varianza explicada por otros predictores. En ese sentido, en el ámbito organizacional y académico se sabe que la inteligencia es el mejor predictor del desempeño (Postlethwaite, 2011; Salgado, 2017), por lo que, si los factores de personalidad no contribuyen sobre las medidas de inteligencia en la predicción de cualquier variable criterio relativa al contexto organizacional o académico su utilidad práctica y teórica sería cuestionada.

Salgado y De Fruyt (2005), al examinar la validez predictiva añadida de los Big Five sobre el desempeño ocupacional y académico, encontraron que los factores conciencia, estabilidad emocional y amigabilidad añadían un 30.30%, 9.07% y 20.12% de validez, respectivamente, sobre la medida de inteligencia en la predicción de criterios ocupacionales. Por su parte, respecto al criterio éxito en la formación, hallaron que conciencia (24.20%), apertura a la experiencia (22.24%), extraversión (18.04%) y amigabilidad (7.47%) añadían validez sobre las medidas cognitivas. Salgado y De Fruyt (2005) concluyeron, por tanto, que una medida de inteligencia junto con un test de personalidad basado en los Big Five serían una óptima combinación para predecir el desempeño. Resultados que se han refutado en otros estudios, consolidando a utilidad de los Big Five como medida predictiva del desempeño.

Sin embargo, a pesar de la evidencia encontrada sobre las medidas de personalidad, la inquietud sobre su uso no se ha disipado. Numerosos autores han manifestado su preocupación sobre la vulnerabilidad que presentan las medidas de personalidad frente al fenómeno del faking en la selección y evaluación de personal (Viswesvaran y Ones, 1999).

En el capítulo anterior se ha analizado la prevalencia del faking en los contextos de SP y hemos comprobado que las personas, bajo las circunstancias apropiadas, tal y como plantean los modelos teóricos descritos, pueden falsificar sus respuestas para obtener un mejor resultado durante su evaluación. Por ello, parece indiscutible la necesidad de conocer cuáles son los efectos que puede llegar a causar el faking para poder desarrollar herramientas específicas que atajen las consecuencias que este fenómeno pueda provocar.

A continuación, se presenta una revisión de los principales trabajos empíricos que han investigado los efectos del faking sobre las medidas de personalidad, primordialmente, y los procesos de evaluación para, seguidamente, exponer los modelos teóricos que han surgido para explicar los efectos encontrados.

### **3. EFECTOS DEL FAKING SOBRE LAS MEDIDAS DE PERSONALIDAD**

La evidencia empírica ha mostrado resultados dispares a la hora de describir cómo actúa el faking y qué consecuencias reales provoca sobre las medidas no cognitivas y los procesos de selección (Donovan et al., 2014; Lebreton, Barksdale, Robin y James, 2007). Como resultado, se puede encontrar en la literatura un diverso conjunto de estudios primarios, meta-análisis y revisiones que han tratado de esclarecer cuánto de perjudicial es la presencia del faking (Holden y Book, 2012; Salgado, 2016; Tett, Anderson, Ho, Yang, Huang y Hanvongse, 2006; Ziegler et al., 2012).

La mayoría de estas investigaciones han estudiado los efectos que tiene el faking centrándose en las medidas de personalidad por ser, como hemos indicado, uno de los instrumentos más extensamente utilizados en los procesos de contratación y han tratado de explicar las consecuencias que tiene sobre cuatro aspectos esenciales: (a) las puntuaciones de los tes; (b) la fiabilidad de las medidas; (c) la validez de las medidas y (d) el ranking de los candidatos. En las próximas líneas se resumen los principales resultados que se han hallado para cada uno de estos aspectos.

#### **3.1 EFECTOS DEL FAKING SOBRE LAS PUNTUACIONES**

Comprender cómo influye el faking sobre las puntuaciones de los candidatos en las pruebas ha sido el principal objetivo de la investigación sobre faking. La mayoría de los estudios que han analizado esta cuestión se han basado en la comparación de las



medias entre personas que respondían a las pruebas en condiciones honestas y bajo condiciones que incitaban a cometer faking, tanto en contextos reales como experimentales.

En su mayor parte estos estudios primarios han encontrado que la distorsión provoca un aumento (o disminución cuando se trata de faking negativo) de las medias de las puntuaciones en los factores de personalidad (Hough et al.; 1990; Douglas et al., 1996). Estos resultados de estudios primarios se han visto ratificados en varios estudios meta-analíticos (Viswesvaran y Ones, 1999; Birkeland, et al., 2006; Hooper, 2007; Salgado, 2016).

El meta-análisis llevado a cabo por Viswesvaran y Ones (1999) se ha considerado como una de las mayores evidencias empírica de los efectos del faking sobre las medidas de personalidad tradicionales (*single stimulus*; de aquí en adelante SS) Estos autores examinaron la capacidad de los sujetos para cometer faking y las implicaciones que tiene sobre las medidas de personalidad. En su trabajo, compuesto por 51 estudios experimentales, analizaron de modo independiente los resultados para muestras inter-grupos e intra-grupo. En el caso de los diseños inter-grupos, los resultados mostraron que los participantes aumentan sus puntuaciones en torno a media desviación típica, los valores  $d$  estandarizados fueron  $d = 0.64$  para estabilidad emocional,  $d = 0.63$  para extraversión,  $d = 0.65$  para apertura a la experiencia,  $d = 0.48$  para amigabilidad y  $d = 0.60$  para conciencia. En las muestras intra-grupo, encontraron un mayor impacto del faking sobre las medidas con valores promedio de  $d = 0.70$ , siendo los valores del tamaño del efecto para estabilidad emocional  $d = 0.93$ , para extraversión  $d = 0.54$ , para apertura a la experiencia  $d = 0.76$ , para amigabilidad  $d = 0.47$  y para conciencia  $d = 0.89$ . Este meta-análisis también demostró que el faking actúa disminuyendo la desviación típica de las puntuaciones. Por tanto, concluyeron que este fenómeno provoca un aumento de las medias de las puntuaciones en los test y una disminución de los valores de la desviación típica. Ambos procesos unidos producen una homogeneización artificial de las muestras.

Paralelamente, otros autores se han interesado por estudiar los efectos del faking sobre las puntuaciones en entornos reales de selección, como un indicador del desempeño típico, es decir, la diferencias que se dan de modo natural en un contexto laboral normal. Entre los estudios primarios que más se han citado se encuentran los trabajos de Rosse et al. (1998), Hough (1998a) o Stark, Chernyshenko, Chan, Lee y Drasgow (2001). Sus investigaciones muestran que los candidatos obtenían puntuaciones medias significativamente más altas que los trabajadores.

No obstante, fueron Birkeland et al. (2006) quienes llevaron a cabo un meta-análisis con el objetivo de corroborar los resultados obtenidos en los estudios primarios y determinar si el faking producía los mismos efectos en el ámbito laboral real que los encontrados en estudios experimentales. Este meta-análisis incluyó 33 estudios independientes que comparaban muestras de candidatos y de trabajadores. Los resultados apoyaron el supuesto del que parten todas las hipótesis relativas al faking, los solicitantes de trabajo distorsionan sus respuestas en procesos de selección reales para tener más posibilidades de ser contratados. Además, se comprobó que obtienen puntuaciones promedio más altas que los ocupantes en cuatro de los factores de personalidad, siendo los valores más altos para los factores conciencia ( $d = 0.45$ ) y estabilidad emocional ( $d = 0.44$ ). Los tamaños del efecto estimados para los restantes factores de personalidad fueron  $d = 0.11$  para extraversión,  $d = 0.16$  para amigabilidad, y  $d = 0.13$  para apertura a la experiencia. Al comparar estos datos con los encontrados por Viswesvaran y Ones (1999) en su meta-análisis, se observa que los tamaños del efecto obtenidos en contextos reales siguen la misma tendencia de los encontrados en los estudios de laboratorio, aunque sus valores son ligeramente inferiores. Birkeland et al. (2006) indicaron que una posible explicación de estas diferencias podría ser que en la condición experimental el faking es inducido, es decir, se les pide explícitamente a los participantes que distorsionen sus respuestas y este hecho puede que resulte en un ligero aumento de las puntuaciones respecto a las que obtendrían en contextos reales.

Hooper (2007) refutaría en su meta-análisis los resultados encontrados en los meta-análisis previos al concluir que el faking produce un aumento de las medias en las puntuaciones de personalidad, especialmente en los factores conciencia y estabilidad emocional, y que estas diferencias son sustancialmente más elevadas en estudios realizados en contextos experimentales. Sus resultados se resumen en la Tabla 4.



Tabla 4. Resultados del Meta-análisis de Hooper (2007) que Compara los Efectos del Faking entre Estudios Experimentales y Correlacionales para los Big Five

	<i>K</i>	<i>N</i>	<i>d</i>	<i>SD<sub>d</sub></i>	<i>IC 90%</i>	
					Inf.	Sup.
Experimentales						
EE	24	26,086	0.75	.35	0.64	0.87
EX	26	50,460	0.42	.29	0.32	0.51
AP	16	15,850	0.23	.32	0.10	0.36
A	18	32,675	0.52	.23	0.43	0.61
C	27	60,684	0.75	.42	0.61	0.88
Correlacionales						
EE	98	12,709	0.93	.59	0.84	1.03
EX	99	11,922	0.63	.33	0.58	0.69
AP	46	7,038	0.37	.29	0.30	0.44
A	57	8,797	0.57	.39	0.49	0.66
C	75	13,387	1.07	.47	0.99	1.16

*Nota.* EE = estabilidad emocional; EX = extraversión; AP = apertura a la experiencia; A = amigabilidad; C = conciencia; Experimentales = estudios en contextos de laboratorio; Correlacionales = estudios en contextos reales; *K* = número de tamaños del efecto; *N* = muestra acumulada; *d* = tamaño del efecto ponderado por tamaño de la muestra; *SD<sub>d</sub>* = desviación estándar del tamaño del efecto ponderado; *IC 90% Inf. / Sup.* = valores inferior y superior del intervalo de confianza del 90%. Tabla adaptada de “Self-Presentation on personality measures in lab and field settings: A meta-analysis”, de A. C. Hooper, 2007. *Tesis doctoral*, Universidad de Minnesota pág. 122.

El número de publicaciones que han investigado esta cuestión no ha parado de crecer en los últimos años. Algunos estudios han discrepado sobre el alcance real de los efectos del faking sobre las puntuaciones. Por ejemplo, Ellingson et al. (2007) y Hogan, et al. (2007) argumentan que las diferencias en las puntuaciones apenas son notorias y es poco probable que afecten al proceso de toma de decisiones. Pero la mayoría de estos trabajos obtiene resultados que, de nuevo, demuestran el importante cambio que se produce en las puntuaciones como consecuencia del faking (p. ej., Boss, König y Mechels, 2015; Donovan et al., 2014; Jeong, et al., 2017; Krammer, 2017; Peterson, Griffith y Converse, 2009).

Entre ellos, cabe resaltar el meta-análisis recientemente publicado por Salgado (2016) que ha contribuido a incrementar el apoyo empírico a los efectos que tiene el faking sobre las puntuaciones de los tests de personalidad. Como se puede observar en la Tabla 5, Salgado (2016) compara los efectos de faking en muestras de candidatos, de trabajadores y en diseños intra-grupo. Destacan los valores del tamaño del efecto de los factores estabilidad emocional, extraversión y conciencia, cuya magnitud es superior a la de los restantes factores en las tres muestras. Asimismo, si hacemos una comparación de los resultados de la muestra de trabajadores y de la de candidatos, estos últimos tienen en

promedio tamaños del efecto más elevados. Además, estos hallazgos siguen la línea de los encontrados por Viswesvaran y Ones (1999) para la muestra intra-grupo, a excepción del factor apertura a la experiencia cuyo tamaño del efecto es próximo a cero en el presente meta-análisis y con los encontrados por Birkeland et al. (2006), aunque en este último caso, las magnitudes de los tamaños del efecto son superiores en el meta-análisis de Salgado (2016). Asimismo, si nos centramos en los valores de las desviaciones estándar, en términos generales, son más pequeños para las muestras de solicitantes y las muestras intra-grupo que para las muestras de trabajadores. Estos resultados, por tanto, muestran que el faking también afecta a este estadístico disminuyendo sus valores.

Tabla 5. Resultados del Meta-análisis de Salgado (2016) que Compara los Efectos del Faking entre Tres Tipos de Muestras: Candidatos, Empleados y Diseño Intra-grupo

	$K$	$N$	$d$	$SD_d$	$FR$	$IC_d$ 80%	
						Inf.	Sup.
Candidatos							
EE	11	32,599	1.11	.23	1.28	0.81	1.41
EX	11	32,599	0.77	.17	1.13	0.55	1.00
AP	12	32,917	-0.14	.49	0.99	-0.77	0.48
A	10	31,203	0.56	.17	1.07	0.34	0.78
C	13	33,002	1.18	.29	1.17	0.82	1.55
<i>Promedio</i>			<i>0.70</i>	<i>.27</i>	<i>1.13</i>		
Empleados							
EE	18	5,467	0.67	.34	1.19	0.24	1.10
EX	20	5,798	0.84	.48	1.14	0.22	1.47
AP	20	5,645	0.38	.38	1.06	-0.10	0.87
A	18	5,467	-0.19	.30	0.98	-0.58	0.20
C	21	5,896	0.45	.27	1.07	0.11	0.80
<i>Promedio</i>			<i>0.43</i>	<i>.36</i>	<i>1.09</i>		
Intra-grupo							
EE	4	427	1.98	.24	1.69	1.66	2.29
EX	5	607	0.50	.21	1.07	0.23	0.79
AP	5	607	0.09	.02	1.01	0.07	0.11
A	5	607	0.94	.55	1.13	0.24	1.64
C	5	607	1.52	.13	1.24	1.36	1.68
<i>Promedio</i>			<i>1.01</i>	<i>.23</i>	<i>1.23</i>		

*Nota.* EE = estabilidad emocional; EX = extraversión; AP = apertura a la experiencia; A = amigabilidad; C = conciencia; *K* = número de tamaños del efecto; *N* = muestra acumulada; *d* = tamaño del efecto ponderado por tamaño de la muestra; *SD<sub>d</sub>* = desviación estándar del tamaño del efecto; *FR*: ratio de faking; *IC<sub>d</sub>* 80% Inf. / Sup. = valores inferior y superior del intervalo de credibilidad del 80%. Tabla adaptada de “A theoretical model of psychometric effects of faking on assessment procedures: Empirical findings and implications for personality at work” de J.F. Salgado, 2016, *International Journal of Selection and Assessment*, 24, pág. 218.

Por último, Salgado (2016) llevó a cabo otros dos meta-análisis para analizar los efectos del faking sobre las puntuaciones de personalidad. En el primero de ellos comparó dos submuestras de trabajadores, diferenciando entre trabajadores motivados a cometer faking y trabajadores no motivados. De nuevo, los resultados reflejan un remarcable aumento del tamaño del efecto en la muestra motivada con una  $d$  promedio de .51 frente a .12 para el caso de la no motivada. En el segundo de los meta-análisis estudió el tamaño del efecto a nivel de las facetas asociadas a cada factor. Los hallazgos vuelven a secundar los resultados descritos en la muestra general. La muestra de candidatos obtiene una  $d$  promedio ( $d = 0.60$ ) muy superior a la encontrada en la muestra de trabajadores ( $d = 0.19$ ). En relación con las desviaciones típicas, de nuevo, se demuestra que el faking reduce los valores de este estadístico en las muestras motivadas a mentir ( $SD = .27$ ) frente a las no motivadas ( $SD = .36$ ).

En resumen, todos los hallazgos meta-analíticos apoyan empíricamente que el faking provoca una distorsión de la puntuación en los factores de personalidad que se traduce en un aumento artificial de la media de las puntuaciones, especialmente en los factores conciencia y estabilidad emocional, y en una disminución de las desviaciones típicas dando lugar a una homogeneización artificial de las muestras.

La evidencia empírica de que los sujetos modifican los resultados de personalidad sugiere una segunda pregunta: *¿lo hacen siguiendo el mismo patrón de respuesta en todos los casos o varía en función del puesto?* Como hemos visto, los modelos teóricos sobre los antecedentes del faking consideran que las variables personales y contextuales influyen sobre los sujetos a la hora de tomar la decisión de cometer faking e incluso se ha planteado que influyen sobre el modo en el que acaban distorsionando sus respuestas. Una parte de la literatura ha tratado de responder a esta cuestión. Birkeland et al. (2006) en su meta-análisis encontraron que el tipo de trabajo al que se opte y el formato del cuestionario influyen en el grado de distorsión de los solicitantes. Así, según las características del puesto, los sujetos tienden a centrarse más en distorsionar ciertos factores, aunque en la mayoría de los casos estos son, de nuevo, conciencia y estabilidad emocional.

Estos resultados apoyan las conclusiones de otros autores como Furnham (1990), Mahar, Cologon y Duck (1995) o Pauls y Crost (2005b). Furnham (1990) realizó un estudio primario con una muestra de 55 personas para identificar si existía un patrón específico de respuesta de faking. En este estudio se les pidió a los participantes que distorsionasen sus respuestas en base a un puesto específico de trabajo (banquero, librero

y ejecutivo) y comparó las respuestas con un grupo control. Concluyó que los sujetos que tienen mayor capacidad para falsificar llegan incluso a modificar perfiles muy específicos. Además, estos hallazgos exponen la dificultad de controlar el faking, al existir una amplia variedad de respuestas distorsionadas, resulta complejo obtener un patrón único de respuestas que se pueda detectar.

Por su parte, Mahar et al., (1995) diseñaron un estudio en el que los sujetos completaban el test de personalidad *Myers-Briggs Type Indicator* (MBTI) en cuatro condiciones diferentes: honestamente, exaltando las características que le harían alcanzar el trabajo, mostrando la mejor versión de sí mismos e imaginando lo que un ocupante del puesto haría. De nuevo, los resultados indicaron que los sujetos cuando quieren distorsionar sus respuestas se basan en los estereotipos que tienen de cómo sería el trabajador ideal para ese puesto. Pauls y Crost (2005b) llegaron a las mismas conclusiones al comparar el efecto que tiene plantear instrucciones generales para faking positivo e instrucciones de falseamiento enfocadas en un puesto específico. Estos autores encontraron que, en la condición general, las puntuaciones aumentaron en todos los Big Five, mientras que bajo las instrucciones específicas sólo aumentaron las puntuaciones entre los factores considerados más relevantes para ese puesto. Incluso, se ha demostrado que los sujetos no siguen el mismo proceso para responder a los test de personalidad, aunque compartan las mismas condiciones experimentales (Zickar et al., 2004).

En conjunto, estas investigaciones han ratificado que las personas, si lo desean o son instruidas para ello, pueden falsificar sus respuestas y provocar un aumento de las puntuaciones que obtienen en los test de personalidad cuando se presentan a un proceso de selección. Además, se ha comprobado que los candidatos tienden a cometer faking en mayor medida que los ocupantes de los puestos. Por último, diversos estudios han concluido que los sujetos no siguen el mismo patrón de comportamiento cuando cometen faking, sino que adaptan sus respuestas a las circunstancias que rodean el proceso de selección. Estos resultados reflejan los planteamientos de las teorías de los antecedentes del faking (McFarland y Ryan, 2000; Snell, Sydell y Lueke, 1999, entre otras), si las conductas del faking están afectadas por el conjunto de variables personales y contextuales que rodean al sujeto, es totalmente justificable que el patrón de respuesta deseable varíe entre los evaluados (Zickar y Sliter, 2012).

### 3.2 EFECTOS SOBRE LA FIABILIDAD

El concepto de fiabilidad hace referencia a la parte de la medida que está libre de error (Dipboye, 1992) y se puede definir como *“el grado en que las medidas nos dan los mismos valores cuando se utilizan repetidamente”* (Schmitt y Klimoski, 1991, pág. 89). De este modo, la fiabilidad, al igual que la validez, nos permite establecer los criterios de calidad de las medidas utilizadas en los procesos de selección (Schmitt y Klimoski, 1991). En ese sentido, cuando la fiabilidad de una medida decrece, también disminuye la calidad de la medida para estudiar los constructos deseados.

Se han desarrollado diferentes estimadores para calcular la fiabilidad de las medidas y la principal diferencia entre ellos es el grado de error que logran corregir. Se pueden diferenciar cuatro posibles fuentes de error cuando medimos las diferencias individuales a través de las medidas de autoinforme (Hunter y Schmidt, 2004): (a) el error aleatorio, que se produce debido a cambios a puntuales (momentáneos) en la atención, eficiencia mental o, por ejemplo, debido a distracciones (Salgado, 2015); (b) el error transitorio, es el error que varía ante diferentes ocasiones pero es constante dentro de cada ocasión; (c) el error específico del factor, debido a las interpretaciones personales que hace cada sujeto de los ítems que conforman la medida y (d) el error debido a diferencias entre jueces. Respecto a los estimadores de la fiabilidad, los coeficientes de consistencia interna, entre los que destacan alfa de Cronbach y el coeficiente de Spearman-Brown, son los más utilizados, aunque sólo permiten estimar el grado de error aleatorio de la medida. El coeficiente test-retest o coeficiente de estabilidad, permite estimar el error aleatorio y el error transitorio. Por su parte, el coeficiente de acuerdo entre jueces estima las variaciones en la medida que pueda deberse a las diferencias individuales entre los evaluadores. Y, por último, el coeficiente de equivalencia y estabilidad (CES), que se obtiene a partir de dos evaluaciones en dos momentos temporales suficientemente separados, es el estimador más preciso de la fiabilidad porque controla las cuatro fuentes de error posibles. Sin embargo, por sus propias características, se necesita tener dos medidas separadas en el tiempo, es menos frecuente que se pueda calcular este coeficiente cuando se estima la fiabilidad de una medida.

El análisis de los efectos del faking sobre la fiabilidad de las medidas de personalidad quizás sea uno de los aspectos que menos interés ha suscitado en la investigación. No obstante, algunos trabajos han aportado datos que nos permiten

examinar los principales efectos encontrados del faking sobre la fiabilidad de las medidas de personalidad.

Stark et al. (2001) estudiaron los efectos del faking a nivel de prueba y de ítem utilizando datos del cuestionario 16PF, obtenidos de dos muestras independientes, una de aspirantes ( $N = 1,135$ ) y otra de personas sin interés en acceder al puesto de trabajo ( $N = 1,023$ ). Tras examinar los efectos de la falsificación sobre la equivalencia de la medida, los resultados mostraron que las fiabilidades ( $\alpha$ ) eran más bajas en la muestra de candidatos, donde oscilaron entre  $\alpha = .59$  y  $\alpha = .85$ , en cambio en la condición honesta ascendían a valores situados entre  $\alpha = .63$  y  $\alpha = .85$ .

Van Iddekinge, Raymark y Roth et al. (2005) encontraron que los coeficientes alfa eran más pequeños en la condición faking que en la condición honesta para las tres escalas que evaluaba el test de personalidad, reduciéndose la fiabilidad promedio de  $\alpha = .82$  a  $\alpha = .73$  cuando respondieron bajo instrucciones de faking. También Heggstad, George y Reeve (2006) hallaron que el faking afecta a la fiabilidad de las medidas de personalidad. No obstante, en su caso encontraron que la fiabilidad de las medidas, cuando se calcula como un coeficiente de equivalencia, estaría sobreestimada debido al efecto del faking. Estos autores desarrollaron un modelo teórico que explicaría estos hallazgos. En esta tesis doctoral explicaremos esta y otras teorías más detalladamente en un apartado específico.

Por su parte, Van Hooft y Born (2012) utilizaron dos medidas independientes de personalidad para estudiar los efectos del faking en su investigación. Todos los participantes cubrieron el *Five Factor Personality Inventory* (FFPI; Hendriks, Hofstee y Raad, 1999) y un test de integridad en condición honesta y de faking. Los estadísticos descriptivos de las condiciones experimentales ponen de manifiesto la atenuación que provoca el faking en la fiabilidad de ambas medidas. En concreto, los valores del coeficiente de consistencia interna o alfa de Cronbach para los factores de estabilidad emocional, extraversión, conciencia y autonomía disminuyeron de .91, .93, .90 y .91 a .84, .84, .85 y .83, respectivamente. Sólo aumentaron, pero muy ligeramente, los valores de alfa para los factores amigabilidad de .81 a .83 e integridad de .71 a .72.

El meta-análisis de Salgado (2016) partía de la hipótesis de que el faking afectaba negativamente los valores de fiabilidad de los Big Five y sus facetas. Los resultados del análisis meta-analítico corroboraron este planteamiento, la fiabilidad de las escalas de personalidad se vio reducida como consecuencia del faking. La fiabilidad promedio era de  $\alpha = .89$  en condiciones honestas (muestras no motivadas a cometer distorsión) y se redujo a  $\alpha = .86$  bajo condiciones de faking. Es decir, la fiabilidad disminuyó en torno a



un 3% cuando se cometía faking. Los resultados del meta-análisis centrados en las facetas de los Big Five mostraron igualmente una reducción de la fiabilidad cuando se comete faking. Si bien en este caso las fiabilidades decrecen a cifras muy por debajo de lo deseable para una herramienta de evaluación de la personalidad, el valor promedio de  $\alpha$  se redujo de .72 a .65.

Estos resultados proporcionan una importante evidencia empírica de los efectos del faking sobre la fiabilidad de las medidas de personalidad y muestran que cuando las personas distorsionan sus respuestas la fiabilidad de las medidas decrece.

### 3.3 EFECTOS DEL FAKING SOBRE LA VALIDEZ

Otra preocupación entre los investigadores y profesionales de SP, son los posibles efectos que el faking provoca sobre la validez de las medidas no cognitivas, especialmente sobre las medidas de personalidad basadas en los Big Five (Hooper, 2007). Se entiende por validez “*el grado en que las inferencias obtenidas de las puntuaciones de las pruebas son correctas*” (Schmitt y Klimoski, 1991, pág. 89). De este modo, un test de personalidad tendrá mayor validez cuando las puntuaciones de los sujetos reflejen de forma más exacta su desempeño. En ese sentido si, como ya hemos mostrado, el faking afecta a las puntuaciones de los test, es predecible que también tenga consecuencias sobre la validez de las medidas. El debate en este contexto se produce a la hora de esclarecer que tipos de validez se verían afectadas y en qué medida (Hough, 1998a).

En su mayoría, las investigaciones que han tratado de esclarecer los efectos del faking sobre la validez de las medidas se han focalizado en dos tipos: (a) la atenuación de la validez de criterio y (b) la modificación de la estructura factorial de las medidas o validez de constructo. A continuación, se presenta un resumen de los principales hallazgos que se han encontrado.

#### 3.3.1 Efectos sobre la validez de criterio

Cuando hablamos de la validez de criterio o validez predictiva nos referimos a la capacidad que tiene la medida de proporcionarnos información para pronosticar una conducta (criterio). En términos estadísticos la validez de criterio es la correlación entre las puntuaciones del constructo medido por la prueba y las puntuaciones de la variable criterio (SIOP, 2003). En la SP, generalmente, la conducta que se desea predecir es el

futuro desempeño de los candidatos para decidir si se adecuan al puesto ofertado y la medida no cognitiva más ampliamente utilizada son los test de personalidad basados en el modelo de los Big Five.

Asimismo, la utilidad de un proceso de selección se basa en la validez de las medidas que se utilicen, si la validez se modifica producirá un impacto directo sobre la rigurosidad de la toma de decisiones (Komar, Brown, Komar y Robie, 2008). Por este motivo, resulta de crítica importancia conocer si el faking produce una disminución de la validez de los predictores de desempeño.

Una visión general de la literatura que ha estudiado los efectos de faking sobre la validez de criterio nos permite realizar dos afirmaciones. La primera de ellas es que existe una gran controversia en cuanto a los verdaderos efectos que producen este tipo de conductas sobre la validez de criterio. La segunda es que, a pesar del cuantioso número de estudios primarios que se han publicado, apenas existen estudios meta-analíticos que aporten conclusiones sólidas sobre esta cuestión.

Ones, Viswesvaran y Schmidt (1993) llevaron a cabo uno de los primeros meta-análisis en que analizaron la validez de criterio de los test de integridad para predecir conductas contraproductivas en el trabajo. En uno de sus supuestos compararon los resultados entre muestras de ocupantes y candidatos. Los valores de validez resultantes (corregidos por falta de fiabilidad en el criterio y restricción en el rango) fueron de  $\rho = 0.40$  para la muestra de candidatos y  $\rho = 0.29$  en el caso de los trabajadores. Estos resultados señalarían que, al menos cuando se utilizan medidas de integridad, el faking no sólo no atenúa la validez, sino que la aumenta (aproximadamente un 38%).

Pocos años después, Ones et al. (1996) publicarían un de los meta-análisis más destacados y citados sobre el impacto del faking (deseabilidad social) en la validez de criterio de las medidas de personalidad basadas en los Big Five. Tras evaluar el faking a través de escalas de deseabilidad social (EDS), encontraron que la validez de las medidas de personalidad no se ve afectada, es decir, no aumenta una vez controlada la distorsión, sino que los valores  $\rho$  permanecen intactos (ver Ones, et al., 1996, pág. 669). Es más, los autores mostraron que lo único que reflejan el faking son las diferencias individuales en los factores de personalidad, al no encontrar una correlación significativa (corregida por falta de fiabilidad en ambas medidas) entre las EDS y el desempeño en el trabajo ( $\rho = .01$ ) pero sí con los factores de personalidad conciencia y estabilidad emocional, con valores  $\rho = .13$  ( $N = 4,850$ ,  $K = 27$ ) y  $\rho = .18$  ( $N = 1,648$ ,  $K = 14$ ), respectivamente. Basándose en estos hallazgos, estos autores concluyen que no sería necesario ejercer un



control específico sobre el faking ya que no se trataría de una medida predictora del desempeño.

Sin embargo, este trabajo ha generado importantes críticas por sus limitaciones conceptuales y metodológicas. Rosse et al. (1998) exponen que Ones et al. (1996) no diferenciaron entre los dos componentes cualitativamente diferentes de la deseabilidad social (ver capítulo 1; Paulhus, 1991). Se sabe que sólo el componente de manejo de impresiones representa las respuestas intencionadas y por lo tanto el único que podría introducir varianza y afectar a la validez y utilidad de la medida. Ones et al. (1996) al no hacer esta diferenciación, pudieron haber obtenido resultados muy sesgados en su meta-análisis. En el aspecto metodológico, Rosse et al. (1998) también apuntan a la presencia de otras posibles problemáticas y sostienen que el análisis de correlaciones puede no verse afectado por los cambios que se pueden producir en el ranking de candidatos cuando se comete distorsión. Aunque la validez observada de la muestra en general no varíe, la validez de los solicitantes que estarían en los primeros puestos del ranking sí podría estar afectada y se aproximaría a cero al tomar en consideración que la distorsión ocurre principalmente entre los solicitantes que reciben las puntuaciones más altas. Además, Rosse y colaboradores (1998) exponen otros factores metodológicos que podrían estar moderando los efectos de la distorsión en la validez predictiva como la distribución sesgada del faking, el ratio de selección, la restricción en el rango y la validez moderada de las pruebas de personalidad utilizadas. Estos autores, en base a estas explicaciones, ponen en duda la generalización de los resultados encontrados por Ones et al (1996).

Para superar las críticas realizadas a Ones et al. (1996), Li y Bagger (2006) replicaron su meta-análisis pero en este caso diferencian entre los dos componentes de la deseabilidad social. Los resultados obtenidos siguieron la misma línea de los encontrados por Ones et al. (1996), tanto el manejo de impresiones como la deseabilidad social apenas afecta a la validez de criterio de las medidas. Al eliminar su efecto no se detectó una atenuación de la validez predictiva de la medida. A pesar de que estos resultados ratifican las evidencias señaladas por Ones et al. (1996), se deben tomar con cautela ya que exclusivamente se han centrado en estudios que han utilizado la medida *Balanced Inventory of Desirable Responding* (BIDR; Paulhus, 1984) y en muchos casos los análisis se han basado en un número reducido de tamaños del efecto.

Contrariamente, Hough (1998a) realizó revisión de la literatura empírica publicada sobre esta cuestión y concluyó que la mayoría de los resultados encontrados en estudios con muestras reales apoyaron la hipótesis de que el faking afecta la validez de

criterio, produciendo una importante atenuación de la misma. Y, además, aclaró que esta disminución de la validez de criterio se agrava a medida que incrementa la distorsión en las pruebas.

En ese mismo año, Ones y Viwesvaran (1998) publicaron otra revisión cualitativa de los principales estudios que habían tratado de analizar el impacto del faking sobre las medidas de personalidad basadas en los Big Five y sobre los test de integridad. Sus conclusiones contradijeron a las obtenidas por Hough (1998a). Tras revisar un amplio conjunto de datos, incluidos los meta-análisis antes mencionados, afirman que el faking no tiene impacto sobre la validez de criterio.

Los resultados encontrados en las citadas revisiones reflejan claramente las discrepancias en sus conclusiones sobre los efectos del faking sobre la validez (Donovan et al., 2003). En este sentido, otros estudios han tratado de esclarecer esta relación. Muchas investigaciones han demostrado que el faking atenúa la capacidad predictiva de las medidas de personalidad, tanto en estudios experimentales (Converse, Peterson y Griffith, 2009; Douglas et al., 1996; Komar et al., 2008; Marcus, 2006; Schmitt y Oswald, 2006; Táuriz, 2011), como en estudios correlacionales (Christiansen, Goffin, Johnston y Rothstein, 1994; Donovan et al., 2014; Jeong, Christiansen, Robie, Kung y Kinney, 2017; Peterson, Griffith, Isaacson, O'Connell y Mangos, 2011). Mientras que otros estudios han defendido que los efectos que se pueden encontrar son mínimos y no afectan a la validez de criterio de las medidas (Barrick y Mount, 1996; Hogan et al., 2007; Hough y Ones, 2002).

Ingold, Kleinmann, König, y Melchers (2015) encontraron un efecto positivo de la falsificación en la validez predictiva de una medida de personalidad. Hallaron, en una simulación de un proceso de selección, que las puntuaciones de personalidad de la muestra de candidatos tenían mayor validez predictiva ( $r = .20/.30$ ) que las puntuaciones de la condición honesta ( $r = -.08/.03$ ). Estos resultados se asemejarían a los encontrado por Ones et al. (1993) en su meta-análisis.

Más recientemente, la revisión de Holden y Book (2012) propusieron que estas variaciones pueden estar justificadas en la diversidad de procedimientos aplicados: diferentes contextos de evaluación y características de las muestras (militares, estudiantes, diversidad de puestos de trabajo), variables criterio (desempeño, calificaciones), operativizaciones del faking (comparaciones entre grupos, grupos de prevalencia, EDS) e incluso diferentes métodos de análisis (análisis de covarianza, moderadores, correlaciones). No obstante, estos autores consideran que se ha demostrado,

examinando los resultados conjuntamente, que el faking tiene un sustancial impacto sobre la validez predictiva de las escalas de personalidad, a pesar de que cuando se analiza en contextos reales de selección (o no manipulados), las discrepancias en los resultados aumentan.

Salgado (2016), tratando de superar las críticas a los resultados de los estudios previos por la diversidad de procedimientos utilizados, analizó, en el meta-análisis más reciente sobre esta cuestión, las consecuencias de este fenómeno sobre la validez de las medidas mediante el coeficiente de homogeneidad ( $u$ ). Encontró un valor  $u$  más próximo a 1 en la muestra motivada a cometer faking, lo que se traduce en que los resultados en la muestra de candidatos son más restringidos. Esto datos mostraron, por tanto, que el faking provoca una subestimación de la validez de los Big Five.

Tabla 6. Coeficientes de Homogeneidad del Meta-análisis de Salgado (2016) entre Muestras de Trabajadores Motivados y no Motivados a Cometer Faking

	$K$	$N$	$u$	$SD_u$	$ICu\ 80\%$	
					Inf.	Sup.
Candidatos						
EE	11	32,599	.81	.051	.75	.87
EX	11	32,599	.82	.049	.76	.88
AP	12	32,599	.78	.063	.70	.86
A	10	31,203	.89	.012	.88	.90
C	13	33,002	.80	.061	.72	.88
<i>Promedio</i>			.82			
Empleados						
EE	18	5,566	.91	.069	.82	1.00
EX	20	5,897	.98	.052	.91	1.05
AP	19	5,518	1.02	.111	.88	1.16
A	17	5,240	.97	.121	.82	1.22
C	20	5,769	.96	.106	.82	1.10
<i>Promedio</i>			.97			

*Nota.* EE = estabilidad emocional; EX = extraversión; AP = apertura a la experiencia; A = amigabilidad; C = conciencia;  $K$  = número de tamaños del efecto;  $N$  = muestra acumulada;  $u$  = coeficiente de homogeneidad;  $SD_u$  = desviación estándar de  $u$ ;  $ICu\ 80\%$  Inf. / Sup. = valores inferior y superior del intervalo de confianza del 80%. Tabla adaptada de “A theoretical model of psychometric effects of faking on assessment procedures: Empirical findings and implications for personality at work” de J.F. Salgado, 2016, *International Journal of Selection and Assessment*, 24, pág. 220.

### 3.3.2 Efectos sobre la validez de constructo

La validez de constructo de una medida se define como “*el grado en que ciertos conceptos (psicológicos, económicos, sociológicos) explican el rendimiento en una medida*” (Salgado y Moscoso, 2018, pág. 32). Se trata de la validez de la estructura interna de la prueba. Parte de la literatura sugiere que el faking, no sólo afectaría a las puntuaciones de los test de personalidad y su validez predictiva, sino que apunta a que también alteraría la estructura factorial de estas medidas. De confirmarse, estos resultados pondrían en serias dudas la validez de las medidas de personalidad en los procesos de selección.

Cuando se trata de analizar las consecuencias del faking sobre la validez de constructo de las pruebas es necesario tener una medida de equivalencia de las pruebas. En ese sentido, el método más ampliamente utilizado es el uso combinado de un análisis factorial exploratorio y confirmatorio. Cuando se encuentran diferencias en la estructura factorial entre los grupos analizados se considera una evidencia de que el faking tiene un efecto negativo sobre la validez de constructo de la medida (Stark et al., 2001). No obstante, existen otros métodos que, aunque menos utilizados, también se han aplicado para detectar cómo afecta el faking a la estructura factorial de las medidas no cognitivas. Por ejemplo, el método IRT se centra en el nivel del ítem y trata de estimar si se producen diferencias en el funcionamiento de los ítems a través de los grupos (Stark et al., 2001).

Los resultados empíricos encontrados sobre esta cuestión han sido contradictorios. Se ha encontrado que el faking puede dar lugar al mismo número de factores, es decir, no producir ningún efecto sobre la validez de constructo, pero también aumentar o incluso reducir la estructura factorial de las medidas de personalidad.

Respecto al primero de los casos, uno de los estudios más tempranos y relevantes fue el desarrollado por Michaelis y Eysenck (1971) que analizaron los efectos sobre la estructura factorial del test Eysenck Personality Inventory (EPI). Estos autores estimaron índices de comparación de factores entre condiciones de alta y baja motivación para falsear y encontraron que la estructura factorial se preservaba, a pesar de la presencia de distorsión en las puntuaciones.

Ellingson et al. (2001) llegaron a la misma conclusión en un estudio más amplio en el que evaluaron cómo afectaba el faking a la estructura factorial en cuatro grandes muestras, provenientes de varias organizaciones, que respondieron a diferentes tests de personalidad muy utilizados en los contextos organizacionales: Assessment of

Background and Life Experiences (ABLE;  $N > 10,000$ ), California Psychological Inventory (CPI;  $N > 6,000$ ), Sixteen Personality Factor (16PF;  $N > 13,000$ ) y Hogan Personality Inventory-Revised (HPI-R;  $N > 2,000$ ). Los tamaños de las muestras oscilaron entre 1,000 y 7,570 sujetos para cada condición y en todos los casos se compararon las respuestas entre sujetos que respondieron honestamente al test y aquellos que manipularon sus respuestas para mostrar una imagen más socialmente aceptable. Tras realizar los análisis factoriales correspondientes y comparar la estructura factorial en las cuatro muestras, los resultados mostraron, en todos los casos, que las estructuras factoriales eran equivalentes entre los grupos, lo que sugiere que la validez de constructo no se ve alterada por las conductas de faking.

Smith, Hanges y Dickson (2001) obtuvieron resultados similares con tres muestras independientes de estudiantes, trabajadores y solicitantes de trabajo, compuestas por 2,500 personas cada una, que contestaron al HPI. Smith et al. (2001) encontraron que la estructura factorial de los Big Five se mantuvo estable en las tres muestras evaluadas. Los estadísticos obtenidos muestran un ajuste adecuado de los datos al modelo en todos los casos con valores para los índices GFI y AGFI superiores a .94 en todos los casos y valores RMSEA entre .06 y .05.

Smith y Ellingson (2002) tampoco encontraron diferencias sustantivas en la estructura factorial en su revisión del HPI entre una muestra de estudiantes ( $N = 406$ ) y otra de candidatos ( $N = 370$ ), los índices de ajuste de los modelos para cada muestra mostraron que las relaciones entre los factores de personalidad no cambiaron sustancialmente al pasar de un contexto motivador ( $X^2 = 336.29$ ;  $p < .005$ ; GFI = .97; AGFI = .96; CFI = .98; RMSEA = .07), a uno no motivador ( $X^2 = 355.60$ ;  $p < .005$ ; GFI = .95; AGFI = .94; CFI = .97; RMSEA = .08).

Los resultados de Marshall, De Fruyt, Rolland y Bagby (2005) fueron muy similares. No hallaron diferencias significativas en la validez de constructo al comparar muestras de candidatos motivados y no motivados a cometer faking en dos estudios independientes, uno realizado en Francia y otro en Bélgica. La congruencia en la estructura factorial entre las muestras motivadas y no motivadas fue del 96% para el primer estudio y 97% para la muestra belga.

Contrariamente a los hallazgos de los estudios citados, otras investigaciones mostraron que la presencia de faking produce modificaciones en la estructura factorial, encontrándose dos tipos de efectos: (a) la reducción de la estructura factorial de las medidas o (b) el aumento de número de factores establecido.

Dentro del primer caso, Ellingson, Sackett y Hough (1999) comprobaron en un diseño intra-grupo con 245 militares, que completaron la prueba Assessment of Background and Life Experiences (ABLE; Hough, Kamp y Ashworth, 1993) que la distorsión daba lugar a la reducción de la estructura factorial de esta medida de personalidad de diez a un único factor. Encontraron que las correlaciones entre las escalas en la condición faking eran extremadamente altas y significativas en prácticamente todos los casos (los valores  $d$  oscilaron entre .58 y .91), resultados que sugirieron la presencia de un único factor en esta condición. Al aumentar las correlaciones entre las dimensiones se produce una reducción en el número de factores extraídos.

Van Iddekinge, et al. (2005) llegaron a resultados similares, aunque en su caso la reducción factorial no resultó tan drástica. Concretamente, estudiaron los efectos del faking en la validez de constructo en una prueba de autoinforme y una entrevista de selección que medían los mismos constructos de personalidad. El análisis factorial mostró que en ambas pruebas el faking afectaba a la estructura factorial pero que el efecto más grave se producía sobre el test de personalidad. Los índices de ajuste obtenidos sugirieron que un modelo de dos factores, en lugar de los tres factores en el que se basa el test, podría explicar mejor los datos obtenidos en la condición faking.

Pauls y Crost (2005b) también encontraron que las instrucciones específicas para cometer faking positivo atenúan la validez de constructo y que los cambios que se producen sobre la validez dependen en gran medida del tipo de faking (faking positivo, faking específico para un puesto de trabajo, etc.) que cometan los participantes.

Por otro lado, algunas investigaciones han concluido con sus resultados que el faking produce más factores que los predeterminados en las medidas. Por ejemplo, Schmit y Ryan (1993) compararon la estructura factorial de NEO-FFI (versión corta del NEO-PI-R; Costa y McCrae, 1992) en una muestra de estudiantes y una de candidatos. Los resultados del análisis factorial sugirieron un claro ajuste de los datos en la primera muestra ( $X^2(80, N = 590) = 200.28; p < .001; GDI = .94; NNFI = .90; NCFI = .96$ ), mientras que en el caso de los candidatos esto no se produce, los valores de los tres índices de ajuste fueron inferiores a .90 ( $X^2(130, N = 590) = 437.79; p < .001; GDI = .88; NNFI = .83; NCFI = .88$ ). Schmit y Ryan (1993) justifican este hecho debido a que, en esta condición, los candidatos tienen una motivación, la de ser contratados, que hace emerger un factor general al que llaman “factor del empleado ideal”, es decir, atribuyen este desajuste al efecto que tiene el faking sobre la validez de constructo y que produce la



irrupción de un nuevo factor general que incluye facetas de cuatro de los cinco factores de los Big Five.

De igual modo, Cellar, Miller, Doverspike y Klawnsky (1996) encontraron que, en una muestra de asistentes de vuelo ( $N = 423$ ) que cubrieron dos pruebas diferentes de personalidad, se producía un mejor ajuste de los datos a un modelo de personalidad de seis factores que al tradicional modelo de cinco (Big Five) cuando se producía faking. Los valores de índice de bondad fueron más próximos a .90 y los valores RMSEA rondaron o incluso fueron inferiores a .05 en el modelo de seis factores, resultados que mostraron el mejor ajuste de este modelo a los resultados en la condición faking (Tabla 7).

Tabla 7. Coeficientes del Análisis Confirmatorio de la Estructura Factorial del Estudio de Cellar, Miller, Doverspike y Klawnsky (1996)

Muestra	$X^2$	CFI	RMSEA
NEO-PI			
<i>Cinco factores</i>	1,059.57	.79	.10
<i>Seis factores</i>	737.79	.85	.06
Goldberg's Test			
<i>Cinco factores</i>	226.99	.91	.05
<i>Seis factores</i>	139.76	.95	.03

*Nota.* CFI = índice de bondad de ajuste comparativo; RMSEA = raíz del residuo cuadrático promedio de aproximación.

El conjunto de resultados de estos estudios primarios muestra que existe un amplio abanico de posibles efectos que provoca el faking sobre la validez de constructo. Esta diversidad de resultados suscita dudas sobre si verdaderamente el faking resulta tan peligroso para la estructura factorial de las medidas, de personalidad primordialmente, como se había planteado. Una de las hipótesis que han surgido para explicar estas discrepancias es que al combinar en una misma muestra candidatos para diferentes puestos puede producir una atenuación los efectos del faking que podrían ser más graves si las muestras fuesen más homogéneas (Holden y Book, 2012).

Salgado (2016) trata de despejar esta incógnita y ofrece una explicación para esta diversidad de hallazgos basándose en los resultados que obtiene en su estudio meta-analítico. Salgado (2016) concluye que el faking aumenta las puntuaciones (ver Tabla 5) y atenúa la validez de criterio de las medidas de personalidad (ver Tabla 6). Estos dos efectos unidos tienen implicaciones sobre la validez de constructo de las medidas de

personalidad. Salgado (2016) aclara que lo que se debe tener en cuenta es que el efecto del faking puede ser univariado o secuencial dependiendo de si es una o son varias las variables distorsionadas, ya que, como se ha concluido anteriormente, no existen un único patrón de respuesta de faking (Birkeland et al., 2006). Si el efecto es univariado se traduce en una subestimación de las correlaciones y, consecuentemente, en una reducción de la estructura factorial. Por el contrario, si el faking es secuencial, produce factores adicionales.

### 3.4 EFECTOS SOBRE EL RANKING DE LOS CANDIDATOS

Más allá de las consecuencias que tiene el faking sobre las propiedades psicométricas de las pruebas existe una cuestión de crítica importancia para los procesos de evaluación y es cómo afecta el faking a la toma de decisiones. Los resultados descritos en las páginas precedentes aportan evidencia empírica de las consecuencias del faking sobre las puntuaciones, la fiabilidad y la validez de las medidas de personalidad, pero estas conclusiones no clarifican si el propio proceso de selección se vería afectado. Algunos autores han destacado esta situación, por ejemplo, Rosse et al. (1998) afirmaron que el faking podía alterar el orden de los candidatos, contratándose a personas menos cualificadas que han inflado sus respuestas para mejorar sus características.

Christiansen, et al. (1994) encontraron discrepancias en el ranking de candidatos en función de si se usaba el test 16PF con o sin una medida de corrección del faking. Concretamente, aquellas personas que fueron seleccionadas tras la valoración con una medida no corregida no lo serían si se controlase el efecto del faking. La discrepancia en el ranking de candidatos, en función del ratio de selección, oscilaría entre un 4%, para un ratio de .95 y un 16%, para un ratio de .15. A la misma conclusión llegarían Mueller-Hanson, Heggstad y Thornton III (2003), hallaron que el porcentaje de personas contratadas que cometen faking aumenta rápidamente a medida que disminuye el ratio de selección. Otros estudios han llegado a resultados similares (Douglas et al., 1996; Zickar, 2000; Zickar, Rosse y Levin, 1996), cuando se produce faking el listado de seleccionados sufre importantes cambios respecto a la condición honesta.

Alliger y Dwight (2000) indicaron que el análisis de la validez de criterio es insuficiente para evaluar el impacto del faking. En su lugar, propusieron que las consecuencias del faking se determinan por un examen de la calidad de la toma de decisiones en la contratación de los empleados. Además, plantearon que, aunque el faking



no afectase a la validez de la medida, no implicaría que este fenómeno no afectase al proceso de selección. Para justificar su planteamiento expusieron el caso de la selección mediante la estrategia *top-down*, en la que sólo los sujetos que obtengan puntuaciones más altas en las pruebas pasan el proceso de selección. Indicaron que puede darse el caso de que pocas personas hayan distorsionado sus respuestas y esto conlleve que los estadísticos descriptivos no reflejen los efectos que provoca el faking sobre las puntuaciones y la validez. Sin embargo, este comportamiento ha permitido a esos sujetos estar entre los candidatos mejor valorados. Por es muy probable que pasen el corte de selección y sean contratados, a pesar de que su puntuación real en las dimensiones evaluadas era más baja. Por consiguiente, debido a los afectos del faking, incluso aunque un escaso número de personas lo cometan, se estaría contratando a gente no cualificada para el trabajo y, por consiguiente, afectando al proceso de toma de decisiones. Por último, el estudio realizado por Griffith et al. (2007) es uno de los más recientes que tratan esta cuestión. Sus resultados son inequívocos, el faking modifica gravemente el ranking de candidatos en los procesos de selección.

En general, los estudios sobre los cambios que produce el faking en el ranking de los candidatos se focalizan en aquellos procesos de toma de decisiones que utilizan la estrategia *top-down*. Con el uso de esta técnica sólo se contratarían a aquellas personas que están en los puestos más altos del ranking ya que se consideran que son los más cualificados para realizar el trabajo. Sin embargo, las personas al falsificar sus respuestas ocupan estar primeras posiciones no por méritos propios si no por los cambios que han realizado de modo consciente en sus respuestas desplazando a los demás solicitantes y perjudicando la calidad de la selección (Morgeson, Campion, Dipboye, Hollenbeck, Murphy y Schmitt, 2007; Táuriz, 2011)

Por lo tanto, se puede concluir que la falsificación puede atenuar la utilidad de las medidas de personalidad en los procedimientos de contratación, especialmente cuando se aplica la estrategia *top-down*, al cambiar el orden de clasificación de los solicitantes.

A continuación, se detallan las principales conclusiones de este capítulo:

1. Las medidas de personalidad, específicamente las basadas en los Big Five, han demostrado ser una importante herramienta predictiva de relevantes criterios ocupacionales y académicos. Concretamente:

a) En lo que respecta al desempeño ocupacional, la evidencia meta-analítica ha demostrado que todos los factores de personalidad son predictores de

algún criterio relativo al trabajo. No obstante, el factor conciencia es el más relevante por ser el único que generaliza su validez a todos los puestos y ocupaciones (p. ej. Barrick y Mount, 1991; Salgado 1997).

b) En el ámbito académico también el factor conciencia se ha identificado como un importante determinante del desempeño académico y otros criterios relacionados (Poropat, 2009; Salgado y Táuriz, 2014), seguido de los factores estabilidad emocional y amigabilidad.

2. Las medidas de personalidad no están exentas de críticas. Una de las que más preocupación e incertidumbre genera en la actualidad es la importante vulnerabilidad que presentan estas medidas frente al fenómeno del faking.

3. Los estudios sobre los efectos del faking muestran el efecto perjudicial que produce este fenómeno sobre las puntuaciones y propiedades psicométricas de las medidas de personalidad:

a) El efecto sobre la media de las puntuaciones es uno de los más estudiados. Los meta-análisis de Viswesvaran y Ones (1999) en el ámbito experimental y de Birkeland et al. (2006), Hooper (2007) y Salgado (2016) en el ámbito laboral real han mostrado que los evaluados modifican sus medidas de personalidad y llegan a aumentar sus puntuaciones en torno a media desviación típica. Si bien este efecto es más reducido en los estudios en contextos reales, el factor conciencia seguido de estabilidad emocional, son los que sufren un mayor cambio en todos los casos.

b) Asimismo, el faking también afecta a las desviaciones típicas de las puntuaciones, disminuyéndolas (Salgado, 2016). Este efecto, junto con el aumento de las medias, provocan una homogeneización artificial de las muestras hecho que dificultaría la labor de diferenciar qué personas son más idóneas para ser seleccionadas afectando, de este modo, a todo el proceso de evaluación.

c) La fiabilidad también se ve afectada por el faking. Los resultados meta-analíticos demuestran que la fiabilidad decrece significativamente. Es decir, aumenta el grado de error que contiene la medida cuando se comete faking.

d) Los estudios más recientes sobre la validez de criterio han mostrado que el faking provoca una atenuación de la misma. Cuando los sujetos distorsionan sus respuestas se obtiene una medida más imperfecta, por lo que necesariamente implica una reducción del poder predictivo de la misma.

c) En cuanto a la validez de constructo, se evidencian diferentes efectos sobre las medidas dependiendo de si el faking afecta a una (univariado) o varias de las variables evaluadas (secuencial). En el primer caso dará lugar a una reducción de la estructura factorial. Por el contrario, si el faking es secuencial, producirá factores adicionales.

4. Respecto al impacto sobre la toma de decisiones en los procesos de selección, el faking permite que aquellas personas que mejor distorsionan sus respuestas ocupen los primeros puestos en el ranking de contratación, relegando a los sujetos que han respondido honestamente. En ese sentido, el uso de la estrategia *top-down* fomenta la contratación de candidatos que falsean.





## CAPITULO 3

# MODELOS TEÓRICOS EXPLICATIVOS DEL FAKING





Conocer los antecedentes del faking y desarrollar de un modelo teórico que permita explicar sus efectos son dos cuestiones de absoluta relevancia para los procesos de evaluación que facilitarían la creación de herramientas eficaces para su detección y control.

Este capítulo presenta una revisión de las variables y de los principales planteamientos teóricos sobre los antecedentes del faking. Se describen asimismo los modelos más relevantes sobre los efectos y consecuencias de este fenómeno.

## **1. VARIABLES ANTECEDENTES ANALIZADAS EN LA LITERATURA CIENTÍFICA**

El faking es un fenómeno complejo. Existen diferentes formas de cometerlo y, por consiguiente, diferentes conductas que se pueden asociar al mismo. Por ello, es razonable considerar que son múltiples las variables que pueden afectar a su aparición (Griffith et al., 2006).

Con el objetivo de sintetizar las principales características estudiadas como potenciales desencadenantes del faking, se ha llevado a cabo una revisión de las publicaciones que han analizado las conductas de faking y sus antecedentes. Concretamente, se han examinado las revisiones narrativas sobre faking de Heggstad (2012), Him (2011) y Bolino, Kacmar, Turnley y Gilstrap (2008), el trabajo de Tett et al. (2006) que examina a través del planteamiento de siete preguntas la importancia y efectos de este fenómeno sobre las medidas de personalidad y la posterior revisión realizada por Griffith y Robie (2013). Además, se han revisado diversos estudios primarios (Dar-Nimrod, Ganesan y MacCann, 2018; Jonason, Lyons, Baughman y Vernon, 2014; LaHuis, 2005; Pauls y Crost, 2005a; Raymark y Tafero, 2009, entre otros).

Un primer análisis de estos trabajos nos muestra que se han considerado muchos y muy dispares predictores del faking. Varias de estas variables han sido ampliamente analizadas y los estudios las han descrito como importantes antecedentes del faking. Otras variables, sin embargo, no han mostrado relación con el faking. En todo caso, lo que sí se pone de manifiesto es la consideración de que los diferentes comportamientos de faking tienen lugar debido a la interacción entre el contexto y las características personales (Ziegler et al., 2012). Por ello, son estos dos conjuntos de variables las más estudiadas en la literatura (Bott, Snell, Dahling y Smith 2010; Kacmar, Carlson, Bratton, 2004).

### 1.1 CARACTERÍSTICAS INDIVIDUALES O DISPOSICIONALES

Se ha contemplado una gran diversidad de variables individuales (o disposicionales) como desencadenantes de las conductas de faking. Los estudios y revisiones narrativas citadas se han centrado especialmente en los factores de personalidad. Aunque también variables sociodemográficas, las capacidades cognitivas y las percepciones han sido examinadas. En la Tabla 8 se puede visualizar una síntesis de las características individuales examinadas como antecedentes del faking.

Tabla 8. Características Individuales Relacionadas con la Tendencia a Cometer Faking

<i>Categoría</i>	<i>Variables examinadas</i>
1. Variables sociodemográficas	Sexo, edad, raza/etnia, religión.
2. Variables de personalidad	Modelo de los Big Five (estabilidad emocional, extraversión, amigabilidad, apertura a la experiencia y conciencia), integridad, autoconciencia, maquiavelismo, rasgos psicopáticos, narcisismo, ambición, competitividad, locus de control, autocontrol, autoeficacia, manejo de impresiones, descabildas social, frialdad ( <i>coolness</i> ) exageración ( <i>over-claiming</i> ), autoestima, necesidad de aprobación, inteligencia emocional, empatía, imaginación.
3. Variables cognitivas	Capacidad cognitiva, capacidad intelectual, sesgos cognitivos (negación, aquiescencia, sesgo temporal).
4. Percepciones y actitudes	Ética, valores éticos, moralidad, código moral, norma subjetiva, actitudes hacia el faking, percepción, de justicia, percepción de control sobre el comportamiento, percepción de la habilidad de engañar, percepción de la frecuencia con la que se cometer faking, creencias sobre las actitudes y comportamientos de otros.
5. Otras variables	Conductas contraproductivas.

#### 1.1.1 Variables de personalidad

La personalidad es una de las variables que más atención ha acaparado en el estudio de los antecedentes del faking. Se ha considerado que algunos rasgos de personalidad podrían determinar que un individuo cometa faking (p. ej., Griffith et al., 2006; McFarland y Ryan, 2000; Snell, et al. 1999). Estos rasgos influirían sobre las



actitudes, interpretaciones y comportamientos individuales. Estabilidad emocional y conciencia han sido los factores con los que se han encontrado correlaciones más altas (McLarnon, DeLongchamp y Schneider, 2019; Ones et al., 1996).

*Integridad.* Esta variable se relaciona con conceptos como la honestidad o la justicia. Para las personas íntegras disimular su verdadera manera de ser al responder a un test de personalidad sería una actitud incorrecta. En ese sentido, algunos autores han sugerido la existencia de una relación negativa entre este rasgo y las conductas de faking. Por ejemplo, McFarland y Ryan (2000) encontraron que las personas altas en integridad están menos inclinadas a manipular sus respuestas. Es decir, a más integridad, menos probabilidad de cometer faking.

*Maquiavelismo.* También es uno de los rasgos de personalidad más estudiados como antecedente de faking. Aquellas personas altas en esta variable son más propensos a tener conductas de faking (Bolino y Turnley, 2003; Jonason et al., 2014).

*Rasgos psicopáticos.* Book, Holden, Starzyk, Wasylkiw y Edwards (2006) plantearon que las personas que manifiestan rasgos psicopáticos tendrán más facilidad para distorsionar su imagen cuando responden a un test de personalidad. Propusieron esta relación partiendo de la definición de psicopatología de Mealey (1995): “*el uso social del engaño para aprovecharse de otros*”. Book y colaboradores realizaron un estudio con 201 estudiantes con el que obtuvieron apoyo empírico a su hipótesis. Jonason et al. (2014) también obtuvieron resultados similares.

*Narcisismo.* Forma, junto con la psicopatía y el maquiavelismo, la denominada Tríada Oscura de la personalidad, por lo que también se ha considerado como un determinante de las conductas de engaño. Jonason et al. (2014) encontraron una conexión entre la presencia de este rasgo y la estrategia de engañar.

*Locus de control.* Se entiende como el grado en el que el individuo siente que tiene control sobre su vida. Se diferencia entre locus de control interno, en el que el comportamiento y sus consecuencias dependen del propio individuo, y locus de control externo, donde la persona considera que lo que sucede está fuera de su control personal. El locus de control externo es el que más próximo estaría a las conductas de faking. La lógica de esta relación es que las personas con bajo locus de control (locus externo) se sentirán más libres de distorsionar sus respuestas en procesos de selección. La literatura ha apoyado este planteamiento, se ha encontrado que aquellos que son más altos en locus de control interno toman decisiones más éticas y, por lo tanto, son menos propensos a

cometer faking y se siente totalmente responsables de sus comportamientos (Coleman y Mahaffey, 2000; Raymark y Tafero, 2009; Treviño y Youngblood, 1990).

*Autocontrol.* Es la autorregulación del comportamiento. Aquellas personas que puntúan alto en autocontrol tienen un gran dominio de sus conductas, por lo que se ha considerado que son más propensas a cometer faking para mostrar una imagen que se adecue en mayor medida al contexto. McFarland y Ryan (2000) y Wrensen y Biderman (2005) encontraron una relación significativa entre la capacidad de cometer faking y esta variable. LaHuis (2005) Turnley y Bolino (2001), Fandt y Ferris (1990) y Shore y Wayne (1993) también han determinado que esta variable es un importante antecedente de las conductas de faking.

*Autoeficacia.* LaHuis (2005) realizó un estudio en el que analizó el efecto moderador de esta variable focalizándola en la búsqueda de trabajo. Los resultados que obtuvo concluyen que la eficacia que se siente a la hora de encontrar trabajo afecta al comportamiento de los candidatos en los procesos de selección.

Además de las citadas, otras variables de personalidad que se han examinado por su posible relación con la intención de cometer faking han sido la exageración (*overclaiming*) como tendencia de personalidad, la autoestima o la inteligencia emocional, encontrándose son potenciales predictores de las respuestas socialmente deseables (Bing, Klumper, Davison, Taylor y Novicevic, 2011; Marcus, 2009; Mesmer-Magnus et al., 2006).

### 1.1.2 Variables sociodemográficas

*Sexo.* El sexo una variable sociodemográfica que se ha considerado como antecedente del faking. En términos generales, los estudios han mostrado que las mujeres tienden a cometer menos faking que los hombres (Bolino y Turnley, 2003; Graham, Monday, O'Brien y Steffen, 1994; Newstead, Franklyn-Stokes y Armstead, 1996). Además, se ha encontrado que el sexo modera la relación entre las actitudes y la intención de cometer faking (Tibbetts, 1997).

*Edad.* También se ha considerado como un moderador de las actitudes hacia el faking, planteándose que la edad está inversamente relacionada con la realización de conductas de faking. Graham et al. (1996) encontraron que los jóvenes cometen conductas de engaño en mayor medida. No obstante, algunos autores han indicado que estos resultados no pueden considerarse concluyentes (Snell et al., 1999).

*Raza/etnia.* La investigación sobre si la raza es un predictor de las conductas de faking es limitada, siendo el trabajo de Granello y Wheaton (1998) único, hasta el momento, en investigar esta cuestión. Estos autores encontraron una relación predictiva significativa entre ambas variables, concretamente, la población afroamericana tiende a puntuar más alto que la americana.

### 1.1.3 Variables cognitivas

*Capacidad cognitiva.* La literatura ha mostrado que las personas que puntúan alto en capacidad cognitiva son personas que se adaptan más fácilmente al contexto, que tiene una mayor capacidad de aprendizaje y que fácilmente aplican diferentes estrategias para resolver problemas (Hunt, 2011; Salgado y Moscoso, 2008). Por ello, se considera que los candidatos altos en capacidad cognitiva tendrán más facilidad para detectar cuándo es beneficioso cometer faking y, sobre todo, cómo deben responder de modo deseable para ser contratados (Griffith et al., 2006). Sin embargo, las investigaciones han mostrado resultados ambiguos, mientras que algunos estudios no han refutado esta hipótesis (p. ej., Burkhart, Gynther y Christian, 1978; Anderson, Sison y Wester, 1984). Investigaciones más recientes han encontrado una relación significativa entre la capacidad cognitiva y la capacidad de cometer faking (p. ej., Alliger, Lilienfeld y Mitchell, 1996; Mersman y Shultz, 1998; Levashina, Morgenson y Campion, 2009; Wrensen y Biderman, 2005).

*Sesgos cognitivos.* Los sesgos cognitivos, aunque relacionados con la capacidad cognitiva, tienen una particular influencia sobre el faking. Griffith et al. (2006) consideran que estos sesgos afectan a la percepción que tenga el sujeto de sus propios atributos. Los sesgos más comunes que afectan a las respuestas a los test no cognitivos son la negación (tendencia a responder negativamente a los test) la aquiescencia (tendencia a responder siempre afirmativamente) y el sesgo temporal (tendencia a responder de un modo más optimista cuando se describe el desempeño futuro). Este último se explica por la consideración de que las personas tienden a sentir más control sobre su futuro que sobre el pasado o presente y esa seguridad provoca que respondan de un modo más deseable sobre el futuro (Robinson y Ryff, 1999).

### 1.1.4 Percepciones y actitudes

*Valores éticos y moralidad.* Esta variable está relacionada con la variable integridad. Varios autores han recogido en sus planteamientos que la ética, los valores, la

moralidad son importantes determinantes del faking. Boyce (2005) subrayó que la visión de la ética de cada candidato condicionará si comete o no faking y, en la misma línea, Grieve y McSwiggan (2014) pusieron atención en el papel que juegan las normas morales y la ética en la intención de falsear los test psicológicos. Estos autores encontraron que las normas morales personales predicen la intención de cometer faking, una visión positiva de la moralidad del faking incrementa la intención de cometer este tipo de conductas.

*Actitudes hacia el faking.* McFarland y Ryan (2006) y Snell et al. (1999) también plantearon la influencia de las actitudes hacia el faking como un elemento condicionante de su aparición. Buehl y Melchers (2017) analizaron esta hipótesis y encontraron que la actitud hacia el faking es un predictor directo del faking, si las personas tienen una actitud positiva hasta este tipo de conductas serán más propensas a cometerlas.

*Percepción de justicia.* También se ha planteado que la percepción de justicia del proceso de selección y de las pruebas utilizadas afectará al deseo de obtener el trabajo, considerándose esta variable como un moderador de la aparición de conductas de faking (Snell et al. 1999). En ese sentido, Chapman y Webster (2006) encontraron que la percepción que los candidatos tengan del proceso de selección es un determinante de su actitud durante el proceso y por tanto de la elección de mostrar una mejor imagen para alcanzar el trabajo.

*Percepción de control sobre el comportamiento.* Mueller-Hanson, Heggstad y Thornton III (2006) y Pauls and Crost (2005a) propusieron que el control que perciba el candidato sobre la situación y su comportamiento afectará a la intención de cometer faking. A mayor percepción de control más probabilidad de que se cometa faking.

*Percepción de la habilidad de engañar.* Varios autores han planteado que cuánto más seguros se muestren los candidatos de que pueden modificar sus respuestas más propensos se mostrarán a cometer faking (McLarnon et al., 2019; Raymark y Tafero, 2009).

#### 1.1.5 Otras características personales

*Conductas contraproductivas.* McLarnon et al. (2019) han mostrado que aquellas personas que son más propensas a tener un desempeño negativo en el trabajo tienen más probabilidades de distorsionar sus respuestas a los test de personalidad.

## 1.2 CARACTERÍSTICAS SITUACIONALES O CONTEXTUALES

Además de las variables personales, las características situacionales (o contextuales) ha sido el otro gran conjunto de atributos que se han examinado como determinantes del faking. Los factores relacionados con el entorno pueden dificultar o facilitar que se produzcan conductas de faking (Garry, 1953; Kim, 2011). Las revisiones y estudios enumerados al principio de este capítulo también recogen las principales variables situacionales analizadas como determinantes del faking.

Estas variables se han clasificado en dos grandes categorías. La primera engloba las variables situacionales más estrechamente relacionadas con el sujeto y su experiencia y la segunda agrupa las características relativas a las pruebas y procedimientos. La clasificación de estas variables se puede encontrar en la Tabla 9.

Tabla 9. Características Situacionales Relacionadas con la Tendencia a Cometer Faking

<i>Categoría</i>	<i>Variables examinadas</i>
Variables relacionadas con el sujeto	Experiencia del sujeto en otros procedimientos, entrenamiento, percepción del conocimiento sobre el puesto, conocimiento de los requisitos del trabajo, importancia del resultado, percepción de la necesidad de cometer faking, percepción de las consecuencias negativas, expectativas de éxito, valencia, instrumentalidad, expectativa.
Variables relacionadas con las pruebas y procedimientos	Oportunidad de mentir, características de las pruebas (tipo de prueba, formato y modo de puntuación), advertencia, condiciones del proceso de selección, instrucciones de faking.

### 1.2.1 Variables relacionadas con el sujeto

*Experiencia del sujeto.* Se ha planteado que el hecho de tener información previa o haber pasado por situaciones similares afecta a la intención de los candidatos de distorsionar sus respuestas (Levashina y Campion 2007; Snell et al., 1999).

*Valencia, instrumentalidad y expectativa.* Estos tres conceptos, basados en la Teoría de la Expectativa (Vroom, 1964), surgen para explicar cómo el contexto influye en las conductas de faking. La valencia hace referencia a cuánto de apropiado es cometer faking y se basa en la necesidad de conseguir aquello que se desea (Shiflett y Cohen,

1982). La instrumentalidad describe el grado de necesidad de cometer faking. Por último, la expectativa refleja la creencia en las propias capacidades para cometer faking (Ellingson, 2012). Han sido varias las investigaciones que han analizado la relación de cada uno de estos constructos con las conductas de faking. Shiflett y Cohen (1982) encuentran que la valencia es un antecedente del faking mucho más importante que las expectativas o la instrumentalidad, las personas estarán más inclinados a cometer faking cuando consideran que así alcanzarán sus objetivos personales y, por lo tanto, que existe una valencia positiva para cometer faking. Por su parte, Bott et al. (2010) encontraron que las variables expectativa e instrumentalidad correlacionan con los comportamientos de faking.

*Importancia del resultado.* Estrechamente relacionado con el concepto de valencia. Cuánto más relevante sea para las personas alcanzar resultados favorables en las pruebas, existirá una mayor probabilidad de que cedan ante el faking como un modo de alcanzar los resultados deseados (Snell et al., 1999).

### *1.2.2 Variables relacionadas con las pruebas y procedimientos*

*Características de las pruebas.* Variables como el formato de los ítems, el sistema de puntuación o el grado de familiaridad que se tenga con los test, pueden afectar a la intención de los candidatos de cometer. Las pruebas objetivas y aquellas en que es más difícil distinguir cuál es la respuesta correcta son más efectivas contra el faking (Salgado, 2005; Snell et al., 1999). Asimismo, las preguntas centradas en comportamientos pasados se han mostrado más resistentes al faking (Levashina y Campion, 2007).

*Advertencia.* Esta variable también se ha considerado como un antecedente del faking. Se ha considerado que advertir a los candidatos sobre las consecuencias de distorsionar las respuestas a las pruebas no cognitivas reduce el faking (Salgado, 2005). Por ejemplo, Dwight y Donovan (2003) han examinado cómo afecta dar instrucciones con diferentes tipos de advertencia antes de responder a varias pruebas no cognitivas (personalidad y autodesempeño) y Kluger y Colella (1993) estudiaron el efecto de la advertencia en ítems de biodata. Los resultados obtenidos muestran que son una medida efectiva contra el faking.



### 1.3 LA MOTIVACIÓN COMO ANTECEDENTE DEL FAKING

Por último, es necesario hacer una mención específica a una última variable, la motivación. Algunos autores la han catalogado como una variable individual desencadenante del faking. Otros, en cambio, ha considerado que se trata de una característica contextual. No obstante, en todos los casos, se ha tratado la motivación como una variable necesaria para que se produzca el faking. Por estos motivos se ha considerado la necesidad de dedicarle un apartado propio.

Leary y Kowalski (1990) fueron de los primeros en considerar que la motivación es uno de los principales antecedentes que da lugar al faking. Ellingson (2012) propone que la motivación es el antecedente esencial para que tenga lugar el faking. Si el faking es un comportamiento voluntario que persigue un fin concreto, las personas tienen que dirigir su atención para responder de modo socialmente aceptable. Esto hace necesario que exista una motivación por parte del sujeto para que cometa faking (Ellingson, 2012). Sin embargo, si esta variable se analiza de modo pragmático se evidencia que el nivel de motivación no es igual para todas las personas ante las mismas circunstancias e incluso varía en una misma persona ante situaciones semejantes. Por lo tanto, se plantea que la motivación es producto de la interacción entre factores ambientales e individuales (Pinder, 1998) jugando de este modo un papel moderador entre estas variables y la tendencia a cometer faking. De hecho, como se podrá comprobar en el próximo apartado, al analizar los diferentes modelos teóricos que han surgido para explicar las conductas de faking, la motivación se ha considerado en gran parte de ellos como un producto fruto de la interacción entre las variables disposicionales y las contextuales y no como una simple variable antecedente, teniendo un papel clave en la decisión de los sujetos de cometer faking.

## 2. TEORÍAS SOBRE LOS ANTECEDENTES DEL FAKING

La diversidad de teorías sobre los antecedentes del faking que ha generado un marco teórico extenso desde el que guiar las investigaciones empíricas sobre los comportamientos del faking (Griffith y Peterson, 2011; Goffin y Boyd, 2009, entre otros). No obstante, podemos considerar dos como las principales teorías sobre los antecedentes del faking (Goffin y Boyd, 2009). El trabajo de Snell et al. (1999) plantea una teoría centrada en las diferencias individuales y características contextuales como los factores

desencadenantes de la conducta de faking, y la teoría de McFarland y Ryan (2000) que propone un modelo enfocado en las diferencias individuales en el que enfatizan las posibles interacciones entre las variables antecedentes. Pero no han sido los únicos, Goffin y Boyd (2009), Griffith, Lee, Peterson y Zickar (2011), Marcus (2009) y Mueller-Hanson et al. (2006), entre otros, han descrito otros planteamientos que han surgido como variantes de las primeras teorías y que se consideran también en este apartado.

A continuación, se describen brevemente los principales componentes de cada teoría y se mencionan algunos de las investigaciones empíricas más relevantes que les han dado apoyo.

## 2.1 TEORÍA DE SNELL ET AL. (1999)

Snell et al. (1999) plantearon uno de los modelos de referencia sobre los antecedentes del faking (Mueller-Hanson et al., 2006). Propusieron que el faking tiene dos componentes principales: la motivación para cometer faking y la capacidad para cometer faking. Cada uno de estos elementos está afectado por un conjunto de variables disposicionales (o personales) y situacionales (o contextuales) tal y como se representa en la Figura 1.

Con respecto al primero de los componentes, Snell y colaboradores (1999) indicaron que la *motivación para cometer faking* viene definida por tres tipos de factores: demográficos, disposicionales y las percepciones del sujeto. Dentro de los factores demográficos estos autores se han centrado en dos variables, el sexo y la edad. Los factores disposicionales, se refieren a una serie de rasgos estables: manejo de impresiones, integridad, maquiavelismo, manipulación, delincuencia organizacional, locus de control y moralidad. El modelo considera que estas variables disposicionales son relevantes para conocer qué sujetos estarán más motivados para falsear, pero también indica que el contexto tiene una importante influencia en el comportamiento final del sujeto, es decir, la presencia de estas características personales por sí solas no provoca las conductas de faking, si no que el comportamiento final se verá también afectado por la percepción que tenga el sujeto del contexto.

Por último, el factor perceptual o actitudinal, que los autores consideraron como el principal antecedente del faking, incluye un conjunto de variables relacionadas con la percepción que tendrían los sujetos sobre las actitudes y comportamientos de los demás, su percepción de justicia, su actitud hacia el faking, sus expectativas de éxito y la



importancia del resultado. La percepción que tengan los candidatos de estas seis variables afectará a la motivación para cometer faking. Cuanto más común, aceptable y favorable consideren esta práctica, más proclives serán a incurrir en ella.

El segundo componente del modelo, la *capacidad para cometer faking*, vendría precedida por las características de las pruebas (factores contextuales), la experiencia del sujeto y factores disposicionales. No todos los sujetos que llevan a cabo conductas de faking lo hacen de modo exitoso. En ese sentido, estos autores plantearon que existen un conjunto de variables que permiten discernir cuándo un sujeto tendrá la capacidad de llevar a cabo conductas de engaño con éxito. Este éxito se basaría en la capacidad intelectual de los candidatos (factores disposicionales) y su conocimiento del puesto, de las medidas y de los procedimientos utilizados por las organizaciones (experiencia del sujeto). Además, el tipo de pruebas, el formato y el modo de puntuarlas también influiría sobre la capacidad de los sujetos para cometer conductas de faking (características de las pruebas). Snell et al. (1999) incluyeron esta última variable apoyándose en los resultados de investigaciones previas que mostraron que existen medidas más fácilmente falsables que otras (Christiansen, Edelstein y Fleming, 1998).

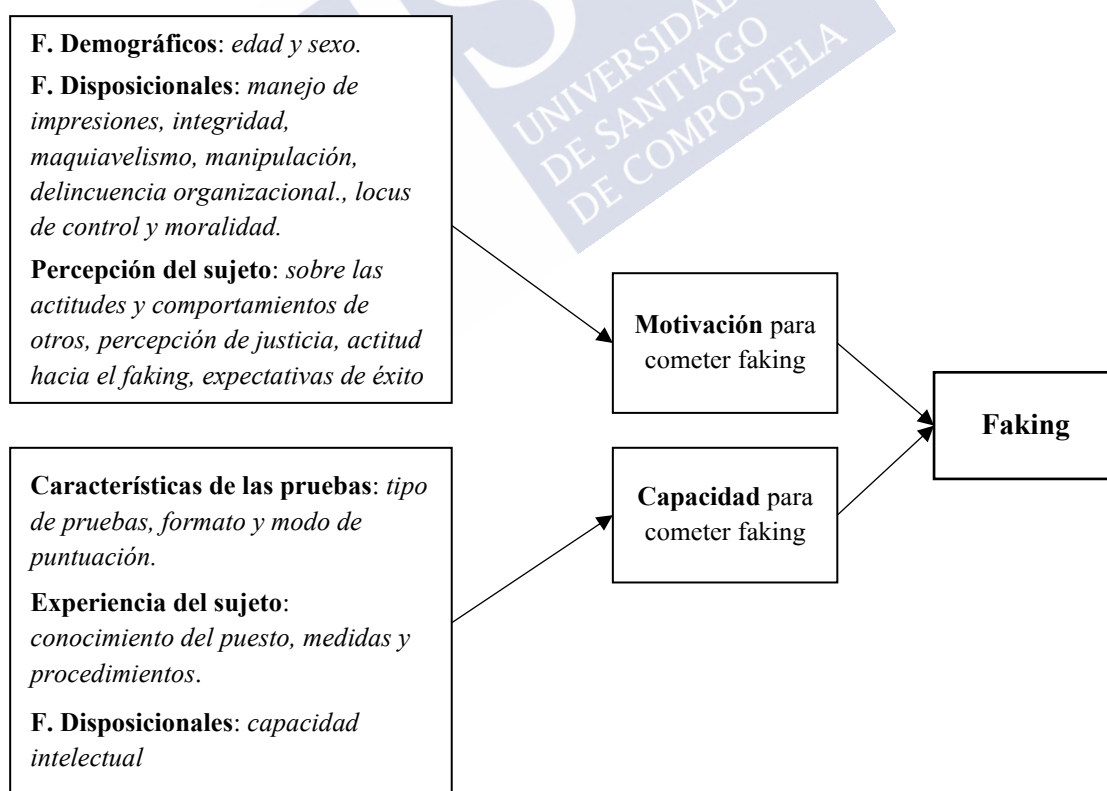


Figura 1. Teoría de Snell et al. (1999). Adaptada de “*Towards a theory of applicant faking: integrating studies of deception*”, de A.F. Snell, E. J. Sydell y S. B. Lucke, 1999, pág. 222.

En resumen, este modelo se centra en esclarecer cuáles son las características personales y situacionales que afectan a lo que los autores consideraron como los dos grandes antecedentes del faking, la motivación y la capacidad para cometer faking, y subraya su importancia para comprender y predecir esta conducta.

Este modelo ha sido parcialmente corroborado en el estudio realizado por Lueke, Snell, Illingworth y Paidas (2001). Estos autores compararon las respuestas de sujetos a un test de personalidad en escenarios de selección hipotéticos y en condiciones honestas y observaron que podían agrupar a los participantes en tres tendencias de respuesta según en nivel de motivación para incurrir en faking. Estos resultados apoyarían la existencia de diferencias personales estables a la hora de cometer faking.

Sin embargo, no se ha profundizado en el estudio específico de los rasgos personales y cómo afectarían a la motivación y capacidad de cometer faking. De hecho, este modelo tiene importantes limitaciones como, por ejemplo, la cuestionable clasificación de algunas de las variables como sólo antecedentes de la motivación y no de la capacidad de distorsionar (p. ej., el sexo o la edad) o el hecho de que sea un modelo apenas estudiado empíricamente (Mueller-Hanson et al., 2006).

## 2.2 TEORÍA DE MCFARLAND Y RYAN (2000)

McFarland y Ryan (2000) se basaron en la teoría planteada por Douglas, McDaniel y Snell (1996) sobre los antecedentes del faking para proponer su modelo. Douglas et al. (1996) plantearon que son tres los factores que dan lugar a las conductas de faking: las características de la situación, las características personales y la oportunidad de cometer faking. Estos autores consideraron que estas variables tienen una relación directa sobre los comportamientos de faking. Sin embargo, McFarland y Ryan (2000) criticaron la simplicidad de este modelo y sugirieron que también existe una interacción entre los antecedentes del faking, es decir, que estas características no sólo afectan directamente a las conductas de faking, sino que también pueden interactuar entre sí y afectar de modo indirecto.

McFarland y Ryan (2000) sugirieron que las conductas de faking vienen determinadas directamente por la intención de cometer faking que, a su vez, estará influenciada por las actitudes que tengan los sujetos hacia este tipo de comportamientos. No obstante, esta relación está moderada por otros dos componentes, la capacidad para

falsear y la oportunidad de llevar a cabo este tipo de conductas. El primero de ellos está determinado por el autocontrol del sujeto y el conocimiento que tenga de las pruebas que va a realizar y el segundo por la capacidad que tenga el sujeto de mejorar los resultados reales que obtendría en la prueba. Es decir, si la puntuación real no le favoreciese en el proceso de selección y poseyera la capacidad para falsear (conocimiento y control de las pruebas) aumentaría su intención de cometer faking y, por lo tanto, sería más probable que llevase a cabo este tipo de conductas. McFarland y Ryan (2000) añadieron una tercera variable de tipo contextual, que englobaría el deseo que tenga el sujeto de alcanzar ese trabajo y las advertencias que tenga durante el proceso sobre las consecuencias de mentir.

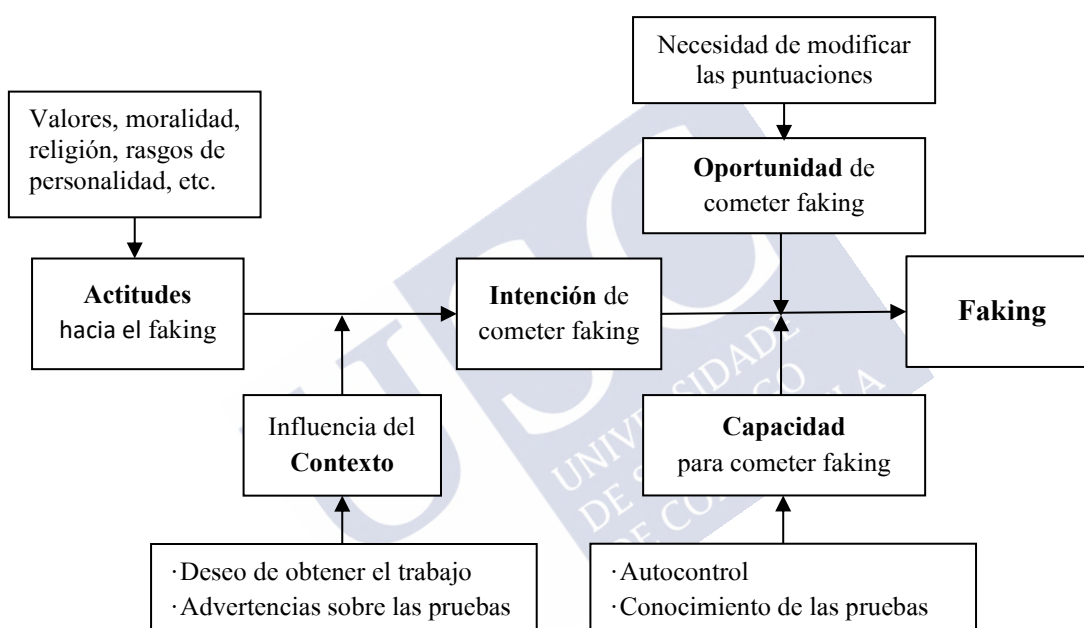


Figura 2. Teoría de McFarland y Ryan (2000). Adaptada de "*Variance in faking across noncognitive measures*", de L. A. McFarland y A. M. Ryan, 2000, pág. 818.

La evidencia empírica ha apoyado parcialmente este planteamiento. Raymark y Tafero (2009) comprobaron que la capacidad para cometer faking está afectada por variables individuales. Concretamente encontraron, en una muestra de estudiantes ( $N = 342$ ) que las variables apertura de ideas (faceta de apertura a la experiencia) y conocimiento de las pruebas correlacionan de modo positivo y significativo,  $\beta = .12$  ( $p < .05$ ) y  $r$  parcial =  $.12$  ( $p < .05$ ) respectivamente, con la capacidad de los sujetos para llevar a cabo conductas de faking. Estos resultados muestran que ciertas variables individuales explican una proporción de la varianza del factor capacidad de cometer faking del modelo de McFarland y Ryan (2000).

Posteriormente, McFarland y Ryan (2006) ampliaron su modelo para integrar la Teoría de la Conducta Planificada (TCP; Ajzen, 1991). Propusieron que el comportamiento de faking se vería afectado por los tres principales componentes de la TCP: las actitudes o “*grado en que la persona hace una evaluación favorable o no favorable de la conducta en cuestión*”, las normas subjetivas definidas como “*la presión social percibida para llevar a cabo o no esa conducta*” y la percepción de control sobre el comportamiento, entendido como el “*grado de facilidad percibida para desarrollar la conducta*” (pág.188). Estos tres componentes, junto a la capacidad de cometer faking y las influencias situacionales, factores de su primer modelo, son las variables que afectan a la intención de cometer de faking. Nuevamente, este planteamiento sólo ha conseguido apoyo parcial y apenas se ha abordado cómo afectarían las variables contextuales a la intención de cometer faking (McFarland y Ryan, 2006; Mueller-Hanson et al., 2006).

Hemos descrito las consideradas como principales teorías sobre los antecedentes del faking. Sin embargo, no han sido las únicas que han intentado dar una explicación plausible a este fenómeno, han surgido otros planteamientos que es necesario mencionar por sus destacables aportaciones teóricas.

### 2.3 TEORÍA DE MUELLER-HANSON ET AL. (2006)

Mueller-Hanson et al. (2006) consideraron que los dos modelos precedentes tenían serias limitaciones que era necesario suplir. Por ello, propusieron un modelo que recoge los aspectos más robustos de los planteamientos anteriores, pero con ciertas variaciones. Su modelo se basa también en que la intención es el precedente inmediato de las conductas de faking y está afectado por la motivación y la capacidad para cometer faking. No obstante, Mueller-Hanson et al. (2006) añaden otros tres factores que afectarían de modo directo a la intención de cometer faking, las percepciones que tienen los sujetos sobre la situación y los factores de personalidad conciencia y estabilidad emocional.

Mueller-Hanson et al. (2006) pusieron a prueba su propio modelo en una muestra de estudiantes universitarios ( $N = 489$ ). Los resultados apoyaron parcialmente su planteamiento. Encontraron una relación positiva entre tres de las variables propuestas como antecedentes y las conductas de faking. Las excepciones fueron la variable capacidad para incurrir en faking con la que no se encontró relación causal y la variable

intención de cometer faking con la que se obtuvo una inesperada relación negativa. Aun así, estos resultados prueban que el faking está parcialmente determinado por la percepción del contexto y por las características individuales de personalidad.

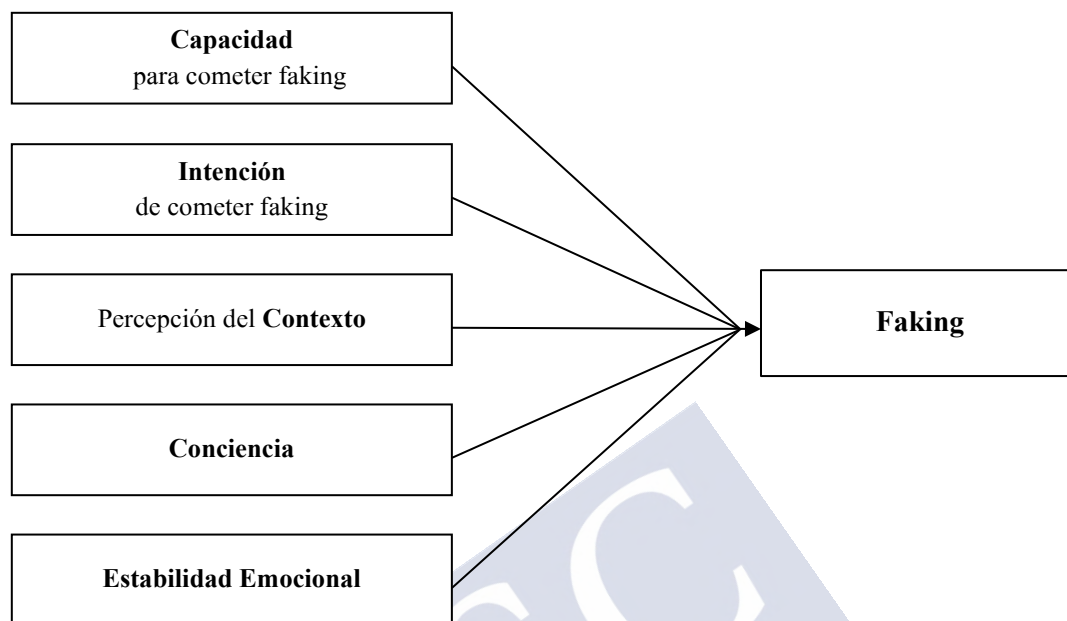


Figura 3. Teoría de Mueller-Hanson et al. (2006). Adaptada de “*Individual differences in impression management: an exploration of the psychological processes underlying faking.*”, de R. A. Mueller-Hanson, E. D. Heggstad y G. C. Thorton III, 2006, pág. 291.

#### 2.4 TEORÍA DE GOFFIN Y BOYD (2009)

Goffin y Boyd (2009) plantearon que para medir y controlar el faking era necesario comprender el proceso psicológico que subyacía al mismo. En ese sentido, estos autores consideraron que los modelos anteriores tenían limitaciones conceptuales que era necesario subsanar. Su modelo nace de los aspectos clave de los modelos previos, sobre los que realizan dos importantes mejoras. Primero cambiaron la perspectiva de análisis del faking. Esta teoría se enfoca en un nivel más específico, la respuesta a los ítems del test, mientras que los modelos previos se centraban en los comportamientos de faking en la prueba en su conjunto. Este cambio en el planteamiento se basa en que los sujetos no tienen respuestas de faking generalizadas durante toda la prueba, sino que realizan una evaluación ítem a ítem de la necesidad de cometer faking.

Respecto a la segunda mejora, Goffin y Boyd (2009) consideraron, al igual que las teorías precededoras (McFaland y Ryan, 2000; Snell et al., 1999), que el faking tiene dos antecedentes principales, la motivación y la capacidad de cometer faking. No

obstante, realizaron una redefinición del segundo componente. Para estos autores la capacidad de cometer *faking* era más subjetiva de lo que se había planteado y la renombraron *percepción de la capacidad de cometer fakeing* haciendo referencia a la percepción subjetiva que tienen los candidatos sobre si conseguirán modificar las puntuaciones para obtener mejores resultados. Consecuentemente, este nuevo enfoque hizo necesaria la revisión de la relación entre las restantes variables propuestas como determinantes del *faking*.

Goffin y Boyd (2009) sugieren que la percepción de la capacidad para cometer *faking*, al ser mucho más subjetiva, no afectaría directamente al comportamiento de *faking*, sino que lo haría de modo indirecto a través de la motivación. El modelo sería más lineal, el comportamiento de *faking* estaría precedido por la motivación para cometer *faking* y éste se vería influenciado por la percepción que tenga cada sujeto de su capacidad de falsear las respuestas. Estos autores añaden que tanto la motivación como la percepción de la capacidad para cometer *faking* estarían afectados por variables de carácter individual y contextual.

Los rasgos de personalidad autocontrol, maquiavelismo, integridad, necesidad de aprobación, conciencia y estabilidad emocional junto con el código moral de cada persona son las variables individuales que afectarían a la motivación para cometer *faking*. La percepción de la necesidad de modificar las respuestas a los ítems y la percepción de las posibles consecuencias negativas son las variables contextuales que afectarían a la motivación para distorsionar las respuestas. En este caso, los autores enfatizaron que, a diferencia de los modelos anteriores, se trata de las percepciones que los sujetos tengan lo que afecta a su motivación, no las consecuencias reales que conllevaría hacerlo.

Con respecto a la percepción de la capacidad para cometer *faking*, Goffin y Boyd (2009) indican que los factores personales que afectarían a este factor también engloban un conjunto de rasgos de personalidad (narcisismo, inteligencia social, innovación, autoestima y conciencia) y la combinación de destrezas, habilidades y experiencia que tenga cada sujeto. En el nivel contextual se incluyen la percepción de tener la oportunidad de cometer *faking* y el conocimiento que tengan los sujetos sobre el puesto.

Con este modelo teórico, Goffin y Boyd (2009) han tratado de concretar la relación entre las principales variables incluidas en modelos pasados (capacidad y motivación para cometer *faking*) y delimitar las diferencias individuales que afectarían a esta conducta.



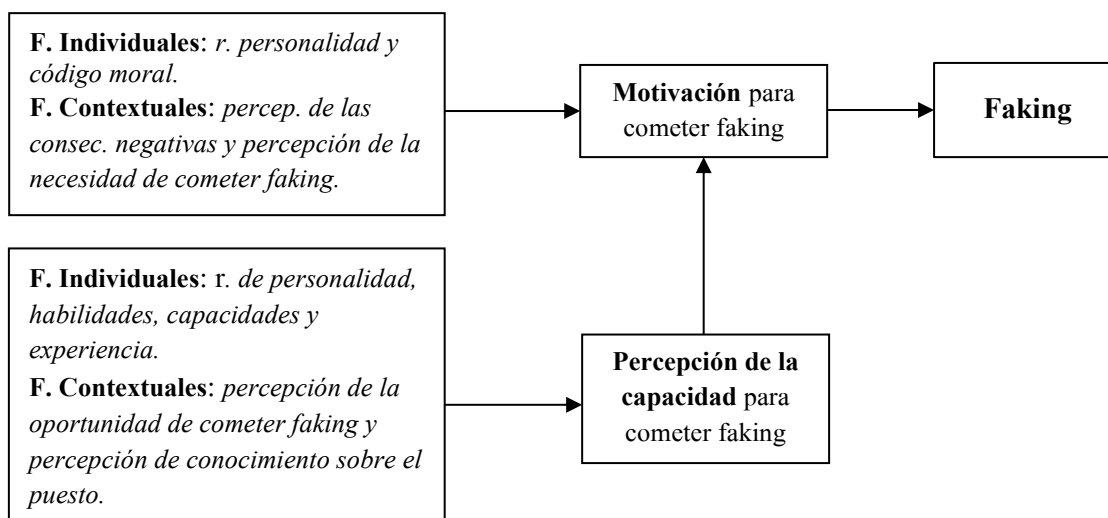


Figura 4. Teoría de Goffin y Boyd (2009). Adaptada de “*Faking and personality assessment in personnel selection: advancing models of faking*”, de R. D. Goffin y A. C. Boyd, 2009, pág. 153.

## 2.5 TEORÍA DE LA CLASIFICACIÓN CONTRATO-RASGO (GRIFFITH ET AL., 2011)

La Teoría de la Clasificación Contrato-Rasgo (*Trait Contract Classification Theory*) de Griffith et al. (2011) también plantea que las características personales juegan un papel fundamental en los comportamientos de faking. Concretamente propone que los rasgos autocontrol, narcisismo, autoestima, impulsividad, (baja) integridad y locus de control (externo) tienen una relación directa con la intención de cometer faking. Es más probable que los sujetos que presentan estas características lleguen a cometer conductas de faking. No obstante, Griffith et al. (2011) incorporaron una novedad respecto a otros planteamientos sugiriendo que el faking es un comportamiento fruto de la interacción entre estas diferencias individuales con una variable de naturaleza contextual, el *contrato psicológico anticipatorio*. Este constructo hace referencia a la relación que el solicitante del puesto percibe que tendrá con la organización, esto significa que los aspirantes tienen en consideración su futura relación con la empresa a la hora de cometer faking.

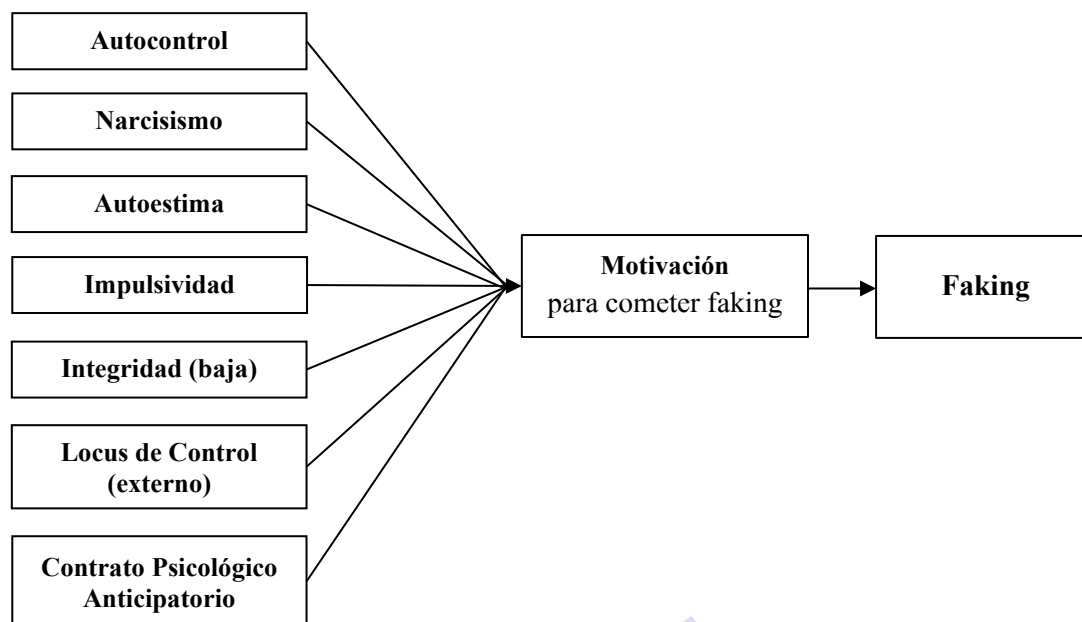


Figura 5. Teoría de la Clasificación. Adaptada de “*First date and little white lies: a trait contract classification theory of applicant faking behavior*”, de R. L. Griffith, L. M. Lee, M. H. Peterson y M. J. Zickar, 2011, pág. 342.

Por último, los autores aclararon que, a medida que los sujetos avancen en el proceso de selección, realizarán una revaluación de la situación que podría afectar a su nivel de motivación para cometer las conductas de engaño, es decir, el nivel de motivación no es fijo ni se tiene en cuenta en un momento único del proceso, sino que podría fluctuar a lo largo de la realización de las pruebas.

Partiendo de esta premisa, los autores consideraron que no existía un único tipo de faking, sino que se pueden diferenciar cuatro niveles que representarían los diferentes procesos de respuesta según el nivel de motivación y las características personales de los sujetos:

1. *Autopresentación (self presentation)*. Es el primer nivel de la taxonomía. Las personas que cometen este tipo de distorsión realzan su imagen, pero siendo fieles a sí mismos. Tienen como objetivo mantener una reputación concreta y con sus respuestas reflejarán una imagen que luego mantendrán en la organización. Se relacionaría con los rasgos de autoestima y autocontrol.

2. *Exageración (exaggeration)*. En este caso los sujetos exageran las respuestas para dar una imagen que excede la realidad. Las respuestas siguen reflejando la personalidad del candidato, pero idealizándolo. Estas respuestas aún se considerarían válidas descripciones de las características del sujeto porque se



basan en sus características reales. Está asociada a los rasgos personales de autoestima y narcisismo.

3. *Respuesta reactiva (reactive responding)*. Es el faking más oportunista. Los sujetos englobados en este nivel son falsificadores “torpes”, impulsivos, que no se basan en su autoesquema para modificar sus respuestas. Contestan a cada ítem sin tener en cuenta las respuestas que han dado previamente. Se vincula a los rasgos impulsividad e integridad.

4. *Respuesta Fraudulenta (fraudulent responding)*. Este último nivel de la taxonomía hace referencia a aquellos sujetos que engañan deliberadamente si tener en cuenta su propio autoconcepto, su único objetivo es alcanzar la puntuación del candidato ideal. Este nivel se relacionaría con los rasgos de locus de control e integridad.

## 2.6 TEORÍA DE MARCUS (2009)

Marcus (2009) presentó una teoría con dos diferencias fundamentales respecto a las teorías que la preceden (McFarland y Ryan, 2000; Snell et al. 1999). La primera, es que adopta una perspectiva diferente a la de modelos anteriores, más centrados en el punto de vista del empleador. En este caso, Marcus (2009) se basó en el punto de vista de los candidatos y sugirió que la mayoría de los solicitantes mienten durante los procesos de evaluación. La segunda diferencia, derivada de la primera, es que el concepto de faking se sustituye por uno más inclusivo, el de *autopresentación*. Este concepto es definido por el autor como “*cualquier intento consciente o inconsciente de controlar las impresiones de los compañeros en las interacciones sociales y que permite describir el comportamiento de los solicitantes al puesto*” (pág. 418). La acepción de este nuevo concepto estaría fundamentada en una serie de evidencias. La primera, el término autopresentación implica que el sujeto valora dos cuestiones a la hora de tomar la decisión de distorsionar sus respuestas, éstas son: ¿cómo consigo el trabajo? y ¿deseo este trabajo? Aunque modelos como el de McFarland y Ryan (2000) sí que incluyeron como variable antecedente el deseo de alcanzar el trabajo, la mayoría sólo ha prestado atención al primero de los interrogantes. La segunda, el concepto de autopresentación no tiene una naturaleza negativa como se asume para el concepto de faking. La tercera, autopresentación es un concepto más amplio que incluye un mayor número de conductas. Y, la cuarta, el término autopresentación va más allá de los contextos típicos de selección.

Asimismo, este autor propuso un modelo sobre cómo sería el proceso que llevaría a un candidato a cometer conductas de engaño durante un proceso de selección. Marcus (2009) aclaró que cada uno de los pasos del proceso de autopresentación estaría afectado por un conjunto de variables situacionales y personales. Las variables situacionales englobarían la necesidad de alcanzar el puesto de trabajo, el conocimiento de las pruebas de evaluación y del puesto, atracción hacia el puesto, y la experiencia. Por su parte, el grupo de las variables personales hace referencia, entre otras, al manejo de impresiones, autocontrol, autoconcepto, empatía, habilidad cognitiva o estabilidad emocional.

De nuevo, este modelo incorporara variables situacionales y personales ya recogidas en modelos precedentes. Sin embargo, este autor especificó que se producirían cambios en la influencia de estas variables a lo largo del proceso de selección. Así, los candidatos comenzarían el proceso de selección con un nivel inicial de motivación para modificar sus respuestas y presentar una imagen más favorable. Pero a medida que avanza el proceso, los solicitantes realizan un análisis continuo de la situación (contexto, pruebas, organización, etc.). Esta nueva evaluación influiría sobre el nivel de motivación del sujeto y determinaría el tipo de autopresentación que realizaría y que se vería reflejado en los resultados de las pruebas. Por lo tanto, el nivel de motivación variaría en función de los resultados del análisis que realizan los candidatos durante el proceso.

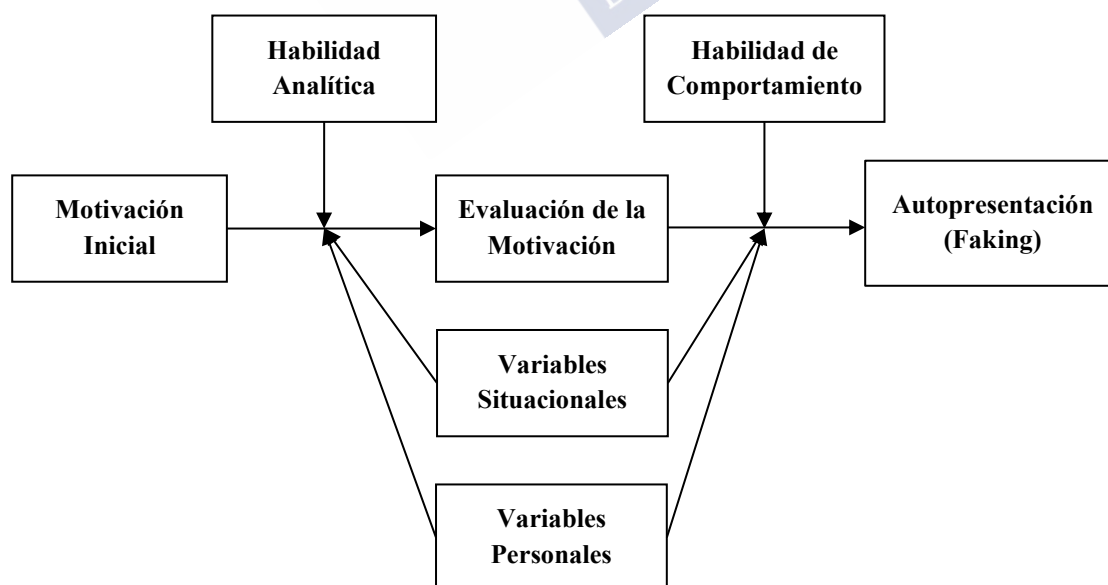


Figura 6. Teoría de Marcus (2009). Adaptada de “*Faking from the applicant’s perspective: a theory of self-presentation in personnel selection settings*” de B. Marcus, 2009, pág. 420.

Paralelamente, han surgido otros planteamientos teóricos que, aunque menos extendidos, han obtenido cierta repercusión en la literatura sobre los determinantes del faking.

## 2.7 MODELO AUTOREGULATORIO DEL FAKING (KOMAR, KOMAR, ROBIE Y TAGGAR, 2010)

Komar y colaboradores (2010) desarrollaron un modelo basándose en la Teoría de la Asignación de Recursos (TAR; Ackerman, 1986; Kanfer y Ackerman, 1989) y en los hallazgos de Robie et al. (2000). La TAR propone que las personas tienen una cantidad de atención limitada que pueden emplear en diferentes actividades relativas al desempeño de tarea, lúdico o a tareas de autocontrol, entre las que se incluirían las conductas de faking. Por su parte, Robie et al. (2000) encontraron que al limitar el tiempo de respuesta a un test de personalidad se reducía la capacidad de los sujetos para poder responder de modo deseable. Estos autores concluyeron que este hecho era debido a que los sujetos tenían que dividir su atención entre responder al test y controlar el tiempo de la prueba. Komar et al. (2010), partiendo de estos hallazgos, proponen que la capacidad de autorregulación y el tiempo dedicado a la prueba son determinantes del faking. Estos autores plantearon que restringir el tiempo reduce la capacidad de los sujetos de asignar recursos cognitivos a la autorregulación, por lo que resulta más complejo que los sujetos distorsionen sus respuestas. Sin embargo, cuando pusieron a prueba su propio modelo con una muestra de 243 estudiantes, los resultados no apoyaron sus planteamientos. Encontraron que no existía una relación estadísticamente significativa entre el tiempo de exposición a la prueba y el faking y tampoco obtuvieron relaciones significativas entre los factores de personalidad y la capacidad cognitiva (autorregulación) cuando se imponía un límite temporal. Estos resultados indicarían que la limitación temporal de las pruebas no supone una exigencia cognitiva para los sujetos y, por lo tanto, no afectaría su intención de distorsionar sus respuestas.

## 2.8 TEORÍA DE LA ACTIVACIÓN DE LOS RASGOS (TETT Y SIMONET, 2011)

Tett y Simonet (2011) propusieron un modelo “*multisaturado*”, la *Teoría de la Activación de los Rasgos*, para explicar las conductas de faking basado en la Teoría Clásica de la Medición (TCM). La TCM plantea que la puntuación observada (X) es la

suma de su puntuación real en el test ( $T$ ) y el error aleatorio ( $e$ ). Por lo tanto, la ecuación se representaría del siguiente modo:

$$X = T + e \quad (1)$$

Tett y Simonet (2011) partiendo de esta simple ecuación propusieron que el *faking* es un tipo de desempeño que afectaría a las puntuaciones de los sujetos y por lo tanto a la puntuación observada del individuo. Por consiguiente, la puntuación observada cuando se comete *faking* resultaría de la suma de la puntuación verdadera del test, del comportamiento de *faking* (*impression management*; *IM*), del nivel de autoengaño (*self-deception*) y del error asociado a la medida ( $e$ ). De este modo la ecuación en este caso sería:

$$X = T + IM + SD + e \quad (2)$$

Explicando el modelo, cuanto más alta sea la puntuación real en el rasgo deseable ( $T$ ) menos intención tendrán los sujetos de cometer conductas de *faking*. Asimismo, cuanto más crean que tienen ese rasgo ( $SD$ ), menos probabilidades tendrán de cometer *faking*. Por último, Tett y Simonett (2011) señalan que la conducta de *faking* depende de tres factores antecedentes ya propuestos en modelos anteriores: la oportunidad de mentir, la capacidad de mentir y la motivación para mentir, cada uno de los cuales está afectado por múltiples influencias individuales y contextuales.

## 2.9 TEORÍA DE LOS HEURÍSTICOS (KUNCHEL, GOLDBERG Y KIGER, 2011)

Los heurísticos son atajos cognitivos que permiten a las personas reformular un problema y simplificarlo para que pueda ser resuelto de un modo más rápido y sencillo (Kahneman, Slovic, Slovic y Tversky, 1982; Tversky y Kahneman, 1986). Según el planteamiento propuesto por Kuncel et al. (2011), los sujetos utilizan un número limitado de heurísticos para tomar la decisión de cometer *faking*. Además, siguiendo a otros autores (Ellingson, 2011, Johnson y Hogan, 2006), plantearon que las metas que desean alcanzar los candidatos son clave para entender cómo responden a los test. Basándose en esta suposición, Kuncel y colaboradores propusieron que el heurístico que se aplique para cometer *faking* dependerá de los objetivos que quieran alcanzar los sujetos y del contexto de selección. Por consiguiente, si se conociese qué heurísticos pueden aplicar los candidatos y ante qué circunstancias los usan, se podría detectar y controlar en mayor medida el *faking*. Sin embargo, estos autores no llegan a profundizar más en este punto y

no aportan una taxonomía de los posibles heurísticos que podrían ser utilizados por los candidatos cuando quieren cometer faking.

## 2.10 TEORÍA VIE: VALENCIA-INSTRUMENTALIDAD-EXPECTATIVA (ELLINGSON Y MCFARLAND, 2011)

Ellingson y McFarland (2011) plantean un modelo basado en la Teoría Valencia-Instrumentalidad-Expectativa sobre la motivación del comportamiento. Sugieren que los candidatos seleccionan una alternativa de acción, cometer o no faking, basándose en las creencias, percepciones y actitudes que tengan sobre cada comportamiento y los resultados que asocian a cada uno. El comportamiento, de este modo, se basa en tres aspectos clave: (a) la valencia o afectividad que cada individuo tiene hacia un resultado; (b) la instrumentalidad, entendida como el grado en que un individuo cree que puede alcanzar los resultados de segundo nivel a través de su desempeño y (c) la expectativa, descrita como la medida en que un individuo cree que conseguirá alcanzar metas de primer nivel. Con resultados de primer nivel, los autores se refieren a aspectos ligados al desempeño, por ejemplo, contestar a un test no cognitivo. Los resultados de segundo nivel son los resultados que se alcanzan fruto del desempeño en el primer nivel, por ejemplo, la oportunidad de ser contratado u obtener un ascenso debido a los resultados del test.

Aplicando este modelo a la conducta de faking, la valencia se refiere al grado en que el individuo desea obtener ese trabajo por encima de otras opciones. La instrumentalidad sería la creencia de que cometer faking ayudará a obtener los resultados adecuados para conseguir el puesto de trabajo y, por último, la expectativa, se refiere a la creencia del propio sujeto de que puede incurrir en faking exitosamente. Además, los autores incluyen un cuarto factor que actuaría de moderador, la capacidad de cometer faking, es decir, la capacidad real que tiene el individuo de llevar a cabo conductas de faking. Cada uno de estos antecedentes del faking, nuevamente, estaría afectado por variables individuales y contextuales.

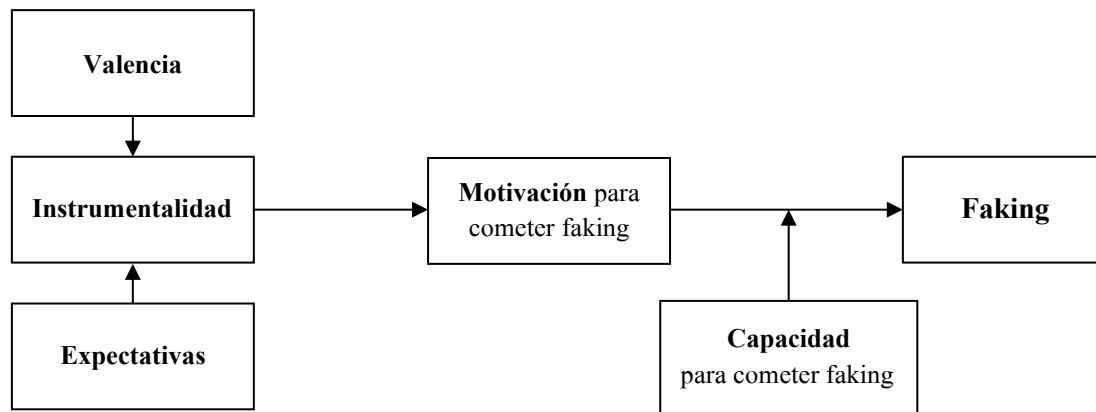


Figura 7. Teoría VIE (2011). Adaptada de “*Undertanding faking behavior through the lens of motivation: an application of VIE theory*”, de J. E. Ellingson y L. A. McFarland, 2011, pág. 326.

### 3. TEORÍAS SOBRE LAS CONSECUENCIAS DEL FAKING

La TCM (Teoría Clásica de la Medición) ha sido el eje central a partir del cual se han desarrollado los modelos estadísticos más importantes que han tratado de explicar los efectos que provoca el faking sobre las medidas de personalidad. Si la TCM define la puntuación que se obtiene en un test como la suma de la puntuación verdadera más un componente de error, el efecto del faking se reflejaría en este segundo componente de error.

En el conjunto de modelos propuestos en base a esta teoría, tres han sido los más destacados por sus relevantes aportaciones para entender los efectos del faking. A continuación, explicaremos los aspectos más importantes de cada uno de ellos, las aportaciones que han realizado sobre el proceso de faking y las consecuencias que provoca sobre las propiedades psicométricas de las medidas.

#### 3.1 TEORÍA DEL FAKING COMO ERROR TRANSITORIO (HEGGESTAD, GEORGE ET AL, 2006)

Heggestad, George et al. (2006) desarrollaron su modelo teórico partiendo de la consideración del faking como error transitorio en las puntuaciones observadas. El error transitorio refleja la interacción entre los sujetos y el contexto de evaluación. Este error varía en cada ocasión, pero es consistente dentro de una misma evaluación, es decir, afecta a todos los ítems medidos en un proceso concreto. El error transitorio puede ser provocado por diferentes variables de tipo contextual tales como el estado de ánimo, una enfermedad temporal o un cambio transitorio en la capacidad de procesamiento de la información.



De acuerdo con estos autores, el faking, al ser error transitorio, reflejaría la varianza en las puntuaciones individuales debido a fuentes no relacionadas con el constructo que se pretende medir sino asociadas al momento de la medición. Por tanto, consideraron que controlar el efecto del error transitorio en las medidas de personalidad es de crítica importancia debido a sus consecuencias sobre la fiabilidad y, por consiguiente, sobre otros parámetros de calidad.

Como hemos expuesto, existen diversos estimadores de la fiabilidad (ver Capítulo 2), pero son los coeficientes de equivalencia (CE) los que más extensamente se han utilizado en el ámbito práctico debido a que para su cálculo sólo se requiere una única administración de la medida. Sin embargo, Heggstad, George et al. (2006) plantearon que, si la varianza de error transitorio es consistente en toda la prueba, al calcular la fiabilidad con un CE se confundiría con varianza verdadera (Schmidt & Hunter, 1999). Consecuentemente, el CE sobreestimaría la fiabilidad del conjunto de puntuaciones que contengan error transitorio y provocaría que los resultados de las pruebas de personalidad se considerasen más fiables de lo que verdaderamente son cuando se ven afectadas por este error. Asimismo, usar los valores de la fiabilidad obtenidos con CE para corregir los coeficientes de la validez de criterio y para calcular el error estándar de medición podría dar lugar a subestimaciones de sus valores reales.

Heggstad, George et al. (2006) propusieron que el uso del coeficiente de equivalencia y estabilidad (CES) permitiría solventar este problema ya que su cálculo permite controlar los efectos debidos al error transitorio. Además, indican que la diferencia entre los valores de CE y CES sería un estimador de la Varianza debida al Error Transitorio (VET). Dado que el faking es error transitorio, los autores señalaron que la estimación de la VET será mucho mayor cuando se comete faking que cuando las puntuaciones son honestas, donde el error sería mucho menor.

Para poner a prueba sus planteamientos, Heggstad, George et al. (2006) realizaron un estudio de medidas repetidas en el que comparan los resultados bajo condiciones de respuesta honesta y faking. El primer hallazgo que encontraron es que bajo condiciones de faking se produce un mayor error transitorio. Este resultado corroboraría su hipótesis de que las conductas de faking son error transitorio y que el contexto influye en el sujeto a la hora de cometer faking. El segundo fue que los coeficientes CE sobreestiman los valores de fiabilidad en condiciones de faking. Para llegar a esta conclusión realizaron una comparativa entre los estimadores CE y CES y concluyeron que el coeficiente CE, en comparación con la medida CES, sobreestima, en

promedio, un 27.5% la fiabilidad en la condición faking, en comparación al 9.7% obtenido para la condición honesta. Por lo tanto, los coeficientes CE no serían una herramienta idónea para estimar la fiabilidad en medidas que han podido ser distorsionadas. En su lugar el uso de CES permitiría un control más efectivo del faking sobre la fiabilidad de las medidas de personalidad.

Si bien, esta teoría propone una explicación estadística del efecto del faking amparada en la evidencia empírica, los propios autores advierten de que se trata de un estudio experimental primario con una muestra de estudiantes, por lo que sería necesario replicar el estudio con otras muestras y en otros contextos para poder generalizar los resultados que encontraron. Asimismo, aunque la aplicación del estimador CES permitiría controlar el efecto del faking, lo cierto es que se trata de un estadístico poco utilizado en la literatura científica por lo complejo que resulta su cálculo al necesitar la administración de dos medidas paralelas en dos momentos diferentes. No obstante, Schmidt, Le y Ilies (2003) publicaron un trabajo con una serie de ecuaciones que permiten obtener el parámetro CES a partir de la aplicación múltiple de la misma medida, lo que facilitaría enormemente su cálculo.

### 3.2 TEORÍA DEL FAKING COMO ERROR SISTEMÁTICO DE LA MEDIDA (ZIEGLER Y BUEHNER, 2009)

Ziegler y Buehner (2009) consideran el faking como un error sistemático de medida o error espurio (Schmidt et al., 2003). En la TCM, el componente de error que contiene la medida se define como un error aleatorio, es decir, que no es debido a influencias sistemáticas. Sin embargo, tal y como propusieron estos autores, el faking se trataría de un error que representaría una varianza sistemática producto de la interacción entre el sujeto y las demandas del contexto. Esto autores asumieron que se trata de un error sistemático debido a que consideraron que el comportamiento de faking no ocurre siempre, pero cuando ocurre lo hace siempre bajo las mismas circunstancias, lo que dificultaría diferenciarlo de la varianza verdadera.

Su modelo surgió de los resultados de estudios meta-analíticos previos que mostraron que, según las características del puesto, los candidatos distorsionan en mayor medida ciertos factores de personalidad (Birkeland et al., 2006). Ziegler y Buener (2009) se basaron en las teorías antecedentes del faking para explicar estos resultados y desarrollar su modelo teórico sobre los efectos del faking. En esta teoría, estos autores



explicaron que el afán de los sujetos por ofrecer una imagen más deseable provoca que modifiquen las puntuaciones en las medidas de personalidad (Birkeland et al., 2006; Hooper, 2007; Viswesvaran y Ones, 1999). Si bien, el faking no afectaría por igual a todos los factores, sino que sólo afectaría a aquellas escalas que fuesen consideradas por los candidatos como más relevantes para el puesto deseado. De este modo, la distorsión estaría influenciada por variables contextuales, tal y como proponen los modelos sobre los antecedentes del faking. Por lo tanto, estos autores llegaron a la conclusión de que la falsificación es específica de contexto y, como tal, puede entenderse como un error sistemático de medida provocado por la interacción entre la persona y las demandas situacionales.

Asimismo, Ziegler y Buehner (2009) propusieron que era posible diferenciar entre los rasgos que han sido falseados y los que no analizando la carga factorial de la medida en faking. El faking, como error espurio, sólo aumentaría las correlaciones entre las variables que han sido distorsionadas, aunque no midan el mismo rasgo o factor. Así, consideraron que el error sistemático de medida es varianza de método común y todos los elementos que han sido falseados no sólo cargarán en su propio factor si no también en un factor de método común que representa el faking. Como el error sólo afecta a algunos de los rasgos, sólo estos rasgos compartirán falsa varianza.

Estos autores realizaron un estudio empírico para apoyar su teoría. Encontraron que el faking modifica la media y la covarianza y, además, comprobaron, a través de un análisis de ecuaciones estructurales, que la varianza provocada por el faking se podía separar de la varianza verdadera de los factores, es decir, corroboraron que el error sistemático, es decir el faking, genera varianza de método común.

### 3.3 MODELO TEÓRICO DE LOS EFECTOS PSICOMÉTRICOS DEL FAKING (SALGADO, 2016)

Salgado (2016) planteó el modelo más reciente sobre los efectos del faking en procesos de evaluación. Este modelo surge con la finalidad de justificar algunos efectos del faking que los anteriores modelos no consideraron en sus planteamientos. Como hemos descrito, las teorías de Heggstad, George et al. (2006) y Zeigler y Buehner (2009) permiten explicar los efectos del faking sobre la fiabilidad o la validez de las medidas. Sin embargo, no explican otros importantes efectos asociados a este fenómeno como son los cambios en el número de factores de personalidad o los efectos del faking sobre la desviación típica. Como las teorías precedentes, para desarrollar su modelo se centró en

los efectos que tiene el faking sobre las medidas de personalidad por su especial relevancia en los procesos evaluativos.

Este modelo se basa en la concepción del faking como un fenómeno de homogeneización de las muestras que reduce el rango de puntuaciones obtenidas por los sujetos, lo que provoca que las personas parezcan más similares entre sí de lo que realmente son. El modelo plantea que el faking opera a través de un doble mecanismo. Por un lado, influye en la media de las puntuaciones de los test de personalidad bien aumentándola (faking positivo) o disminuyéndola (faking negativo). El modelo sugiere que esta influencia se puede calcular como un ratio de faking (RF) dividiendo la media bajo condiciones de faking entre la media de la condición honesta. El valor RF debe ser mayor que 1 si existe faking positivo y menor de 1 si se trata de faking negativo. Es igual a 1 si no se produce distorsión. Este efecto ya ha sido ampliamente establecido en estudios previos, pero no como un valor de RF sino como una  $d$  de Cohen como, por ejemplo, las revisiones empíricas de Birkeland et al. (2006) y Ones y Viswesvaran (1999). El segundo efecto del faking es que reduce la desviación típica en las medidas de personalidad, es decir, afecta a la varianza de las pruebas. Si se comparan las desviaciones típicas de un grupo faking y un grupo honesto el valor que se debe obtener es menor que 1. Este es el primer modelo que examina sistemáticamente este segundo efecto del faking.

En resumen, el faking tiene dos efectos de artefactos simultáneos e importantes, actúa modificando las medias y reduciendo la desviación lo que provoca una restricción en el rango de las puntuaciones obtenidas. La principal consecuencia de este incremento de los artefactos (en faking positivo) es una disminución de la varianza común, que sería más evidente cuanto más extremo sea el faking. Asimismo, esta alteración artificial de la media y la desviación típica producida por el faking tendría consecuencias sobre todos los estadísticos basados en ambas medidas: covarianza, coeficientes de correlación y la validez de constructo.

La covarianza es, esencialmente, el valor de la varianza entre dos variables, nos indica cuánto cambian juntas dos variables y en qué medidas. Su fórmula es la siguiente:

$$COV_x = \frac{\sum(x-\bar{x})(y-\bar{y})}{N-1} \quad (3)$$

En base al modelo planteado por Salgado (2016), el faking tiene dos efectos sobre la covarianza de las medidas. Primero, produce valores más pequeños de covarianza,

debido al incremento en la media del predictor ( $x$ ) cuando se comete faking positivo y, segundo, podría afectar al signo de la covarianza debido a los efectos indirectos del predictor ( $x$ ) sobre la media del criterio ( $y$ ). Dado que  $N$  es constante, cualquier cambio que se produzca en la media de  $x$  cambiará los valores de la covarianza. Estos efectos sobre la covarianza son muy relevantes por las importantes consecuencias que tendrían sobre otros estadísticos que se calculan a partir de este valor como, por ejemplo, el signo del coeficiente de correlación. Por consiguiente, el faking también afectaría al valor de los coeficientes de correlación. Dada la fórmula del coeficiente de correlación:

$$r_{xy} = \frac{COV_{xy}}{\sigma_x \sigma_y} \quad (4)$$

Se puede ver que los efectos sobre la covarianza ( $COV_{xy}$ ) afectarían al signo de la correlación, pero, además, la correlación también depende de la desviación típica y, como ya se ha señalado, este modelo propone que el faking reduce la desviación típica de las puntuaciones. Por lo tanto, la combinación de ambos efectos, valores de covarianza más pequeños y la reducción de las desviaciones estándar ( $\sigma$ ) afectarían directamente disminuyendo la magnitud del valor de la correlación. Esta modificación del valor de la correlación tendría consecuencias negativas sobre la validez de criterio y sobre la estructura factorial.

Respecto a la validez de constructo, el modelo también plantea que se debe tener en cuenta el tipo de restricción en el rango que se produce para conocer las consecuencias reales sobre la validez de constructo. Así, el faking puede producir restricción univariante cuando éste sólo afecta a una variable y restricción multivariante cuando afecta a varias variables. La restricción univariante provoca una disminución de la varianza común y un aumento de la varianza de cada variable con la variable restringida. La restricción multivariante produce factores adicionales que no estarían correlacionados con los factores existentes. Para entender este fenómeno debemos retomar de nuevo los valores de la covarianza y del coeficiente de correlación. Si la covarianza entre las variables afectadas por el faking es mayor y su desviación típica es menor, la correlación entre estas variables será mayor. Al reproducirse este fenómeno entre más variables da lugar a nuevos factores, llamados factores espurios. Thomson y cols. (1938, 1939, 1944) con sus fórmulas matemáticas dieron apoyo teórico a los efectos de la restricción multivariante planteados y recogidos en esta teoría.

Salgado (2016) llevó a cabo un meta-análisis para corroborar los tres principales efectos del faking que plantea en esta teoría: el aumento de la media (faking positivo), la disminución de la desviación típica y los coeficientes de fiabilidad más pequeños. Para ello analizó de modo separado muestras de estudiantes, candidatos y ocupantes de puestos de trabajo. Como ya hemos realizado una descripción más detallada de los resultados obtenidos por Salgado (2016) en los apartados dedicados en el Capítulo 2 a cada uno de los efectos asociados al faking sobre las medidas de personalidad, resumiremos los hallazgos encontrados en las siguientes conclusiones:

1. Los resultados mostraron coeficientes  $d$  y RF superiores a 1 con un promedio de 1.09 para los empleados motivados a mentir, de 1.13 para la muestra de candidatos y de 1.23 para la muestra de estudiantes (diseño intra-grupo), lo que demuestra que los sujetos que cometen faking obtienen puntuaciones más elevadas en los test (ver Tabla 5).

2. Respecto al segundo de los planteamientos, el autor encontró valores  $u$  más pequeños en las condiciones faking, con valores promedio de  $u = .97$  para la muestra de ocupantes del puesto y de  $u = .82$  para la de candidatos. Estos resultados demuestran que se produce restricción en el rango de las puntuaciones, es decir, una disminución de la desviación típica en las condiciones en las que los participantes están más motivados a cometer faking (ver Tabla 6).

3. Los resultados corroboran también el tercer planteamiento. Se encontró que el valor promedio de la fiabilidad fue de  $\alpha = .86$  en las condiciones faking y de  $\alpha = .89$  para las honestas, es decir, era un 3% más pequeño en la condición faking respecto de la honesta, confirmándose que la restricción en el rango produce un decrecimiento en el valor del coeficiente de fiabilidad. En resumen, los resultados encontrados en esta revisión meta-analítica dieron apoyo empírico al modelo teórico presentado por Salgado (2016).

Sin embargo, este autor plantea una importante limitación, este meta-análisis sólo incluyó estudios que hubieran usado como medidas de personalidad los test NEO-PI y NEO-PI-R (Costa y McCrae, 1987,1992) por lo que los resultados encontrados sobre la media y la desviación típica sólo se pueden aplicar a medidas SS y se desconoce si son aplicables a otro tipo de medidas de personalidad como los cuestionarios de EF. Por ello,

el objetivo de esta tesis doctoral es poner a prueba las predicciones de este modelo psicométrico de los efectos del faking en el caso de los cuestionarios de EF.

A continuación, se resumen las conclusiones principales de este capítulo:

1. Existen dos principales variables consideradas como precedentes del faking: la motivación para cometer faking y la capacidad para realizar este tipo de conducta. Aunque la relación puede ser directa o indirecta, todas las teorías los incluyen como los dos principales antecedentes del faking.

2. Esto dos componentes estarían afectados por variables individuales y situacionales.

a) Las variables personales se refieren a un conjunto de características dependientes del individuo que la literatura científica ha asociado estrechamente a las conductas de faking como son las variables de personalidad, principalmente conciencia y estabilidad emocional, otros rasgos de personalidad, como integridad, locus de control, manejo de impresiones o moralidad, variables cognitivas y variables de carácter sociodemográfico.

b) Las variables situacionales engloban aspectos como el conocimiento sobre el puesto, el proceso de selección y las pruebas, la necesidad de conseguir el puesto de trabajo y la percepción de las consecuencias que tenga cometer faking.

3. Las teorías que han surgido sobre los efectos del faking se han basado en la TCM para dar una explicación de las consecuencias que provoca el faking sobre las medidas de personalidad. Si bien se enfocan desde diferentes perspectivas, todas definen el faking como un error de medida que afecta la varianza verdadera de las puntuaciones.

a) La Teoría de Heggstad, George et al. (2006) fue la primera en definir el faking como error de medida. Para estos autores el faking supone un error transitorio en las puntuaciones que explicaría las diferencias entre las medias de los sujetos cuando cometen faking. No obstante, es una teoría demasiado simple que se deja en el tintero una explicación para otras importantes consecuencias asociadas al faking.

b) Ziegler y Buehner (2009) plantearon un modelo que explica la diferencias en las puntuaciones entre los propios ítems del tests. El faking es considerado un error sistemático que depende de la interacción entre el sujeto y las demandas del entorno. Su modelo está altamente influenciado por las teorías sobre los antecedentes del faking.

c) Finalmente, Salgado (2016) planteó la teoría más completa que explica ampliamente los efectos psicométricos del faking a partir de la consideración de este fenómeno como un proceso de homogeneización de las muestras que afecta a la media y la desviación típica y por consiguiente a todos los parámetros que se calculan a partir de ambos.

4. El desarrollo de modelos psicométricos sobre cómo funciona en faking ha permitido entender los efectos que este comportamiento produce sobre las medidas no cognitivas. Sin embargo, todos los planteamientos se han centrado en medidas de personalidad SS, por lo que todavía existe la necesidad de conocer si estos efectos se producen ante otro tipo de medidas de personalidad cada vez más utilizadas en la SP como son los cuestionarios de EF.

# CAPITULO 4

## MÉTODOS DE DETECCIÓN Y CONTROL DEL FAKING







Uno de los principales objetivos de la SP es controlar toda fuente de error que afecte a los resultados de los procesos evaluativos para que, de este modo, la toma de decisiones sea lo más objetiva, fiable y rigurosa posible. Sabiendo que el faking se puede considerar como una de las principales fuentes de error que afecta las medidas evaluativas, el desarrollo de herramientas que permitan detectar este fenómeno es esencial.

Sin embargo, esta tarea no ha resultado en absoluto sencilla debido a la propia complejidad de la conducta de faking. Se han desarrollado y aplicado estrategias de naturaleza diversa a lo largo de la historia reciente de la SP para tratar de atajar el perjudicial impacto que provoca sobre los resultados organizacionales. Si embargo, no todas han obtenido los resultados y el apoyo empírico deseado. A lo largo de este capítulo, se presentan los principales métodos que se han utilizado, sus características, efectividad y las principales críticas que han recibido.

## 1. EVALUACIÓN DEL FAKING

Debido a la importancia de sus efectos, han surgido diferentes métodos para intentar mitigar la respuesta socialmente deseable (Nederhof, 1985; Ones et al., 1996; Paulhus, 1981, 1991). Tal ha sido la producción de herramientas que se han desarrollado que ha dado lugar a diferentes clasificaciones según las características básicas que las definen. Así, Paulhus (1991) distinguió cinco grandes categorías: (1) *técnicas racionales*, aquellas en las que el propio formato del test ayuda a controlar el faking para prevenir la respuesta deseable (p. ej. cuestionarios de EF; (2) *técnicas de análisis factorial*, que se aplicarían durante la construcción del test si el proceso engloba seleccionar los ítems que más carguen en cada factor. Si algunos de los ítems más relevantes representan una opción deseable se eliminaría o se rotaría el factor teniendo en cuenta su deseabilidad. De este modo, se asegura que los restantes ítems muestren las verdaderas diferencias entre sujetos creando un test libre de faking; (3) *técnicas de covariables*, en este caso no se centran en evitar las respuestas socialmente deseables, si no que los test incluyen una escala de detección de la deseabilidad social; (4) *reducción de la demanda*, engloba aquellos métodos que reducen la presión de responder de modo socialmente deseable (p. ej., responder de modo anónimo). Esta categoría engloba métodos diversos como cubrir los test de modo anónimo, usando por ejemplo un alias o código numérico, la técnica llamada *bogus pipeline*, un equipamiento que detecta los comportamientos de faking a través de

parámetros fisiológicos, es una de las técnicas más agresivas. Una alternativa a este método es la inclusión de una escala de deseabilidad social, cuando los sujetos la cubren saben que se está valorando su intención de responder de modo socialmente deseable; (5) *técnicas dirigidas a la minimización del estrés*, se trata simplemente de guías para reducir los niveles de estrés en procesos de evaluación y de este modo controlar aquellos estresores asociados a cometer faking.

Por su parte Ones et al. (1996) distinguieron seis categorías: (1) *persuadir* a los participantes de que no mientan, (2) *advertir* a los candidatos de las consecuencias negativas de mentir en los procesos de selección, (3) utilizar *preguntas clave* para la detección del faking, (4) usar *escalas de deseabilidad social*, (5) aplicar *escalas de corrección* (p. ej. Escala K, del MMPI) y, por último, (6) utilizar *cuestionarios de EF*.

Más recientemente, Robie et al. (2006) diferenciaron siete tipos de métodos que han surgido para tratar de reducir los efectos del faking: (1) *técnicas de EF*, (2) *instrumentos con ítems encubiertos*, (3) *estrategias de advertencia*, (4) *medidas de adecuación*, (5) *tiempos de respuesta* (latencia), (6) *puntuaciones observadas de personalidad* y (7) *medidas indirectas de personalidad*.

Esta diversidad de clasificaciones y nomenclaturas muestra claramente la amplitud de métodos que se han desarrollado para tratar de contener los efectos que el faking puede provocar sobre las medidas. Como comprobaremos, de nuevo, casi la totalidad de las estrategias de reducción del faking están enfocadas a los cuestionarios de personalidad. Si bien, algunas de estas estrategias se pueden aplicar ante otras pruebas no cognitivas.

Para estructurar este capítulo, no nos hemos ceñido a ninguna de las clasificaciones citadas, sino que hemos tomado como referencia las revisiones de Dilchert y Ones (2012), Kuncel y Borneman (2007) y Zickar y Gibby (2006) quienes indican que todas estas técnicas se puede catalogar en dos grandes grupos: (a) *medidas que detectan el faking* y (b) *medidas que reducen el faking*.

## 2. MÉTODOS QUE DETECTAN EL FAKING

Son estrategias que permiten distinguir aquellas personas que están distorsionando sus respuestas de quienes responden honestamente a los test, por lo tanto, se trata de medidas que sólo ayudan a identificar a aquellos sujetos que están mintiendo, pero que en ningún caso reducen la incidencia del faking. Las escalas de deseabilidad

social (EDS), son la medida de detección de faking más extendida, aunque estrategias como los *ítems bogus*, el desarrollo de baremos específicos o el control del tiempo de latencia son otras medidas tradicionalmente utilizadas para identificar si se está cometiendo faking.

## 2.1 ESCALAS DE DETECCIÓN DE DESEABILIDAD SOCIAL (EDS)

Las escalas de detección de la deseabilidad social, más popularmente conocidas como escalas de deseabilidad social tiene su origen en los trabajos de Ruch (1942) y Humm y Wadsworth (1934) considerados como los primeros intentos de desarrollar una prueba que ayudase a determinar qué personas manipulan sus respuestas (Burns y Christiansen; 2006; Zickar y Gibby, 2006). Su desarrollo parte de la concepción del faking como un potencial moderador de la validez de criterio de los cuestionarios de personalidad. Se asume que estas escalas reflejarán las conductas de faking que afectan a la validez de estas medidas (Ones et al., 1996).

Se trata de escalas diseñadas específicamente para medir el grado de falseamiento que contienen las respuestas de los candidatos. Están compuestas por ítems altamente deseables ante los que los sujetos deben indicar su grado de acuerdo, de tal modo que las personas que obtengan puntuaciones altas se considera que están distorsionando sus respuestas para mostrar una imagen más positiva (Salgado, 2005; Taúriz, 2011).

Estas escalas pueden ser pruebas independientes o estar incluidas en los test de personalidad:

*A. Escalas Independientes de Deseabilidad Social.* Paulhus (1991) realizó una revisión de las EDS con mayor repercusión en la práctica profesional de la SP y en la literatura científica sobre la deseabilidad social. Estas escalas han evolucionado de acuerdo con las teorías que han ido surgiendo sobre la deseabilidad social y sus componentes. Por orden cronológico debemos destacar:

1. *Edwards Social Desirability Scale* (EDS; Edwards, 1957). Se ha llegado a considerar como una versión corta del Minnesota Multiphasic Personality Inventory (MMPI; Hathaway y McKinley, 1943). Muy ligada a la evaluación de psicopatología.

2. *Marlowe-Crowne Social Desirability Scale* (MC-SDS; Crowne y Marlowe, 1960). Compuesta por 33 ítems que evalúan comportamientos, actitudes

y acciones típicas. Los sujetos tienen que indicar si cada una de las afirmaciones son ciertas o falsas. Un ejemplo de ítem sería: “*Nunca dudo en obviar mis objetivos para ayudar a alguien en el camino*”.

3. *Children's Social Desirability Scale* (CSDS; Crandall, Crandall y Katkovsky, 1965). Formada por 48 preguntas de verdadero-falso, es la versión para población infantil de la MC-SDS.

4. *Self-Deception Questionnaire* (SDQ; Sackeim y Gur, 1978, 1979). Compuesta por 20 ítems que evalúan sentimientos y pensamientos que sólo el propio sujeto puede reconocer que ha tenido. También desarrollaron la escala *Other-Deception Questionnaire* (ODQ) que evalúa comportamientos visibles que son más difíciles de negar. Ha sido respaldada empíricamente y se considera una de las primeras medidas que más claramente conceptualizó y determinó la presencia de deseabilidad social en las pruebas.

5. *RD-16* (Schuessler, Hittle y Cardascia, 1978). A través de 16 preguntas esta escala analiza el impacto que pueden tener las categorías sociales sobre la actitud hacia el faking.

Estas primeras escalas se basaron en la interpretación de la deseabilidad social como un constructo unidimensional. Sin embargo, tras la consolidación del modelo bidimensional de Paulhus (1991, 2002; ver capítulo 1), surgen otras medidas adaptadas a este nuevo planteamiento. Una de las más consolidadas ha sido el *Balanced Inventory of Desirable Responding* (BIDR; Paulhus, 1984). Este inventario se compone de dos subescalas, de 20 ítems cada una, que evalúan independientemente los dos componentes de la deseabilidad social. Desde su creación se han publicado siete versiones de la escala (Mesmer-Magnus y Viswesvaran, 2006). La más reciente, el BIRD-7 (Paulhus y John, 1998), se compone de tres subescalas que evalúan el manejo de impresiones y dos formas específicas de deseabilidad social: mejora autoengañoso y negación autoengañoso.

*B. Tests que incluyen escalas de deseabilidad social.* Muchos de los cuestionarios de personalidad más importantes tienen incorporados una escala de deseabilidad social (o escala de faking). Goffin y Christiansen (2003) realizaron en su revisión de estas medidas de detección del faking, detallaron las características de las más populares. En la tabla 10 se recoge un breve resumen de éstas.

Tabla 10. Breve Descripción de las Escalas de Deseabilidad Social Incluidas en Cuestionarios de Personalidad

Autores	Cuestionario de Personalidad	Nombre de la EDS	Descripción EDS
Gough (1957)	California Psychological Inventory (CPI)	Good Impression Scale	Deseo de crear una impresión favorable
Comrey (1970)	Comrey Personality Scales (CPS)	Response Bias	Tendencia a la respuesta socialmente deseable (RSD)
Berman (1993)	Employment Inventory	Frankness	Expresión de franqueza y defectos comunes
Inwald, Knatz y Shusman (1983)	Inwald Personality Inventory (IPI)	Guardedness	Negación de carencias y defectos/ tendencia a la RSD
Eysenck (1959)	Maudsley Personality Inventory (MPI)	-	Evalúa la presencia de RSD
Hathaway y McKinley (1943); Meehl y Hathaway (1946)	Minnesota Multiphasic Personality Inventory (MMPI-2)	K Scale	Voluntad de admitir problemas e información personal negativa
		L Scale	Exageración de atributos positivos
Saville, Holdsworth, Nyfield, Cramo y Mabey (1984)	Occupational Personality Questionnaire (OPQ)	Social Desirability	Crear una buena impresión
Jackson (1999)	Personality Research Form (PRF)	Desirability	Mostrar una imagen socialmente deseable
Cattell, Cattell y Cattell, (1968)	16PF (4ª edición)	Faking Good (FG)	Tendencia a la RSD
		Faking Bad (FB)	Tendencia a la RS no deseable.
Cattell, Cattell y Cattell, (1993)	16PF (5ª edición)	Impression Management (IM)	Tendencia a la RSD

*Nota.* Tabla adaptada de “Correcting personality test for faking: a review of popular personality tests and an initial survey of researchers.”, de R. D. Goffin y N. D. Christiansen, 2003. *International Journal of Selection and Assessment*, 11, pág. 342.

Las EDS son una de las herramientas que más tiempo se llevan aplicando para la detección del faking y una de las prácticas más extendida en los procesos de selección (Griffith, Peterson, Quist y Benda, 2008; Zickar y Gibby, 2006). El hecho de tratarse de un método de bajo coste y fácil aplicación y corrección ha favorecido su uso. Por ejemplo,



Goffin y Christiansen (2003) realizaron una encuesta a 36 grupos investigadores de EEUU, Canadá y Europa sobre el uso de este tipo de medidas en los procesos de selección y encontraron que, al menos un 69%, de los encuestados favorecen el uso de herramientas que contienen esta medida correctiva del faking. Aunque, sólo un 56% hace uso asiduo de la EDS.

Aunque a priori, resulten un método adecuado para controlar la incidencia del faking, lo cierto es que son varias las críticas que han recibido por los criterios en lo que se basa para definir quién comete faking. Una de las críticas más importantes que tienen las EDS es que esta medida asocie rasgos que son estables de la persona con conductas de faking, dicho de otro modo, que catalogue como *fakers* a personas que están respondiendo honestamente a la prueba simplemente porque sus características se ajustan al perfil “típico” de distorsionador en el que se basan estas pruebas (Burns y Christiansen, 2006; Salgado, 2005; Swanson y Ones, 2002). De hecho, se ha constatado que el falseamiento solo explica el 10% de la varianza de las escalas de deseabilidad, por lo tanto, que los sujetos obtengan puntuaciones elevadas en este tipo de pruebas no indica que se deba a que estén cometiendo faking, sino que puede tratarse de falsos positivos (Burns y Christiansen, 2011). Esta crítica supone un problema práctico ya que pone en duda la eficacia de este tipo de medidas para detectar el faking objetivamente. Es más, estudios recientes han encontrado que la deseabilidad social que se evalúa en las EDS y el comportamiento de faking no estaban relacionados, identificándolos como constructos independientes, lo que pone más en duda la utilidad de esta medida para reducir el faking (Burns y Christiansen, 2006; Peterson et al., 2008). Una última desventaja es que el uso de esta estrategia también genera ciertos costes de aplicación y corrección, aunque menores que los asociados al uso de otras estrategias, como pueden ser las escalas de EF.

Otras medidas con una metodología muy similar a las EDS, aunque mucho menos extendidas y conocidas, son las Escalas de Respuesta Inconsistente (*Response Inconsistency Scales*) y las Escalas de Infrecuencia (*Infrequency Scales*). La primera medida se centra en analizar la consistencia de las respuestas, de tal modo que cada escala contiene pares de ítems que, aunque miden el mismo constructo, reflejan ideas contrapuestas. De tal modo que si un sujeto contesta lo mismo en ambos estaría entrando en una contradicción a la hora de responder, lo que reflejaría que está cometiendo faking. Este tipo de escalas se caracterizan por componerse de muchos ítems para evaluar de modo fiable la inconsistencia en las respuestas y evitar que se detecten falsos positivos. Además, por norma general suelen estar incluidas en escalas mayores, un ejemplo es la *True Response*



*Inconsistency Scale* del MMPI-2, compuesta de 23 pares de ítems. No obstante, existe el riesgo de que los candidatos contesten a los nuevos ítems teniendo en cuenta la información previa o que puedan revisar sus respuestas para marcar la opción coherente con su primera respuesta. Por su parte las Escalas de Infrecuencia, analizan la frecuencia con que los sujetos escogen opciones de respuesta que reflejan comportamientos poco habituales en la población. La acumulación de este tipo de respuestas es un indicador de faking. Sin embargo, dadas las similitudes con las EDS, presenta la misma limitación, la posibilidad de catalogar como falsificadores a personas que verdaderamente tienen puntuaciones altas en rasgos deseables (Dilchert y Ones, 2012).

## 2.2. TÉCNICA DE LA RESPUESTA EXAGERADA O ÍTEMS TRAMPA (ÍTEMS BOGUS)

Esta metodología se presenta como una alternativa a las escalas de deseabilidad social. Aunque en sus inicios se ha tratado como una técnica para la medición de la capacidad cognitiva (Paulhus y Harms, 2004), también se ha orientado a la evaluación del desempeño, experiencia, habilidades y/o conocimientos de los candidatos. Esta herramienta está diseñada para medir, simultánea pero independientemente, el conocimiento real de los candidatos y las respuestas exageradas sobre sus virtudes. Se trata de una técnica en la que se pregunta por el grado de familiaridad que tiene los sujetos con un conjunto de ítems relativos a un dominio concreto. Entre esos ítems se incluyen tareas o información que, aunque aparentemente relevantes para el trabajo, no se pueden realizar o no existen para el puesto (Paulhus, 2012; Paulhus y Harms, 2004). Estos ítems se administran a los candidatos junto con otras preguntas que incluyen información real relativa al puesto (Kuncel y Borneman, 2007; O'Connell, Kung y Tristan, 2011; Reeder y Ryan, 2012). El grado de exageración del candidato se calcula teniendo en cuenta: (a) la proporción de ítems reales que han marcado como familiares y (b) la proporción de ítems inventados que han marcado como conocidos. Aquellos candidatos que marcan mayor proporción de ítems inventados se asume que están exagerando sus respuestas y, por lo tanto, cometiendo faking. Un ejemplo de ítem de este tipo de medida sería: “*He resuelto disputas mediante análisis isométricos.*” Esta técnica de análisis no existe como tal y, por lo tanto, no se puede aplicar para la resolución de disputas. Si los sujetos la marcada como una tarea familiar se entiende que están exagerando su conocimiento y experiencia.

El estudio de Pannone (1984) aplicó esta técnica y puso de manifiesto su efectividad para detectar a los sujetos que están cometiendo distorsión, aunque no ha sido el único. Anderson, Warner y Spencer (1984) y Dwight y Donovan (2003) también han aplicado este tipo de medidas en sus investigaciones. Asimismo, este procedimiento ha obtenido unos valores de fiabilidad razonables en los estudios en los que se ha utilizado, con valores  $\alpha$  que oscilaron entre .70 y .94. Por lo que se puede considerar como una técnica sencilla y robusta en diferentes contextos: SP, educación e inclusive, marketing (Paulhus, 2012; Roeder y Paulhus, 2009)

No obstante, esta estrategia también tiene sus inconvenientes. Primero, es posible que los evaluados no tengan un conocimiento completo del puesto de trabajo al que optan, especialmente en puestos más complejos o que requieren el uso de herramientas muy específicas y que, por lo tanto, consideren como realistas tales tareas. Es decir, que se catalogue como un comportamiento intencional una respuesta basada en el desconocimiento o la mala interpretación de los ítems, dando lugar a falsos positivos. Asimismo, algunos autores han propuesto que la naturaleza del concepto de “exageración” se asocia al “sentimiento de conocer” es decir, al sesgo de creer que se conoce verdaderamente la información que aportan los ítems falsos. Por otro lado, también se ha analizado si las respuestas pueden estar influidas por un componente motivacional asociado al narcisismo. Esto implicaría que en muchos casos la respuesta exagerada no es intencional si no que tiene un componente cognitivo (sesgo de conocimiento) y motivacional (narcisismo) involuntario (Paulhus, 2012). La segunda desventaja es que se trata de escalas fácilmente manipulables por los sujetos si son conocedores del funcionamiento de este tipo de estrategias y reciben el entrenamiento adecuado (Kuncel y Borneman, 2007).

### 2.3. BAREMOS BASADOS EN MUESTRAS DE SOLICITANTES

Crear baremos específicos para muestras de solicitantes en sustitución de baremos de muestras de población general es otra de las estrategias que se han desarrollado para la detección del faking. Estos baremos se elaborarían a partir de puntuaciones obtenidas en contextos en los que los sujetos están motivados a distorsionar sus respuestas, por ejemplo, procesos de selección o de evaluación de competencias. Estos nuevos baremos tendrían en cuenta que parte de la puntuación normativa, que es común a todos los evaluados, se corresponde con las conductas de faking. De este modo, la baremación

tendrá un mayor ajuste a las condiciones reales de aplicación, al utilizar un marco de referencia que incluye a sujetos que tienen un interés en distorsionar sus respuestas (Salgado, 2005).

Aunque se trate de una de las medidas posiblemente más eficaces para la detección del faking, también tiene inconvenientes, el principal es que su desarrollo es una tarea compleja. La creación de nuevos baremos para una prueba implica necesariamente disponer de medios y recursos suficientes para obtener una muestra altamente representativa de la población que responda al test en condiciones de faking. Por lo tanto, se trata de un proceso laborioso que muchas ocasiones es muy complejo llevar a la práctica profesional.

#### 2.4 TIEMPO DE LATENCIA

Alternativamente, se ha propuesto que el tiempo de latencia puede ser un indicador de la presencia de faking en las medidas (Holden, Kroner, Fekken y Popham, 1992; Paulhus y Holden, 2010; Robie, Brown y Beaty, 2007). Esta aproximación se basa en la idea de que en condiciones de faking los sujetos tienen tiempos de latencia diferentes a cuando responden honestamente. El período de latencia dependería del tipo de faking (positivo, negativo) y de los rasgos sobre los que son preguntados los sujetos (Kuncel y Borneman, 2007).

Centrándonos en el faking positivo, objeto de nuestro interés, el planteamiento más consolidado es que los tiempos de latencia aumentan cuando las personas tratan de mostrar una imagen más positiva de la real. La Teoría Cognitiva de Procesamiento de la Información (TCPI) plantea que a mayor carga cognitiva mayor tiempo se necesita para procesar la información (Jensen, 1985). Aplicando esto a la conducta de faking, se deduce que se requiere un mayor esfuerzo cognitivo para responder de modo deshonesto a un test que cuando se responde honestamente (Fine y Pirak, 2016). Desarrollando este planteamiento, Walczyk, Roper, Seemann y Humphrey (2003) explican que hay que tener en cuentas tres elementos cognitivos cuando los sujetos deciden cometer faking: la respuesta verdadera de los sujetos, el proceso de toma de decisión para elegir si falsifican o no la respuesta y, la respuesta final, que en los casos en los que se comete faking implica un proceso de “reestructuración” de la misma. De este modo, al tener más “pasos” cognitivos, decidir y construir la nueva respuesta, los sujetos tienen un mayor procesamiento mental, lo que se traduce en un mayor tiempo de latencia para presentar

una respuesta. Los resultados de varios estudios apoyan este planteamiento. Por ejemplo, Walczyk, Schwartz, Clifton, Adams, Wei y Zha (2005) hallaron en un contexto experimental que cuando los sujetos mienten tardan más tiempo en responder a los ítems. Fine y Pirak (2016) también encontraron que los tiempos de latencia en la condición faking fueron significativamente más amplios en dos medidas distintas de personalidad. McDaniel y Timm (1990) obtuvieron resultados muy similares, aunque en su caso matizan que la naturaleza de los ítems puede afectar a los tiempos de respuesta. De hecho, se ha investigado la eficacia de esta estrategia a través de diferentes niveles de complejidad de las preguntas, formatos y longitudes y los resultados han mostrado que cuando los candidatos están motivados a falsificar tardan significativamente más tiempo en responder (Brunetti, Schlottmann, Scott y Hollrah, 1998).

Aunque la hipótesis más extendida es que los tiempos son más amplios cuando se comete faking, también ha habido autores que han apuntado a la idea contraria, es decir, que podrían ser significativamente más cortos. Es el caso de Holden et al. (1992). Según estos autores, cuando las personas falsifican sus respuestas lo que hacen es activar un esquema previo que hace referencia a una “imagen idealizada”, como tienen ese esquema activo generan las respuestas “automáticamente” requiriendo aún menos esfuerzo cognitivo que cuando no mienten. Siguiendo esta hipótesis, los sujetos tardarían menos en contestar en la condición faking al reducir el tiempo de procesamiento mental. Holden et al. (1992) encontraron apoyo empírico a su teoría. Los sujetos respondían en menos tiempo cuando reflejaban un esquema congruente. Pero no han sido los únicos, Fekken y Holden (1992) y Hsu, Santelli y Hsu (1989) también hallaron resultados que secundan este planteamiento. No obstante, este posicionamiento es el menos extendido, en cuanto a la detección del faking, basado en esta técnica.

A pesar de la potente base teórica sobre la que se asienta la medición de los tiempos de latencia para detectar el faking, también ha sido criticada. Una de las principales críticas se centra en el punto de vista práctico. Para operativizar el faking utilizando esta técnica es necesario disponer de dispositivos electrónicos que graben los tiempos de respuesta y eso no siempre es posible debido a la limitación de recursos (Brunetti et al., 1998; Dilchert y Ones, 2012). Asimismo, esta técnica no está exenta de catalogar falsos positivos. Además, los tiempos de respuestas también pueden estar afectados por otro tipo de variables contextuales o personales difíciles de controlar a través de esta estrategia. Por último, algunos autores también han considerado que con entrenamiento puede ser un método manipulable (Kuncel y Borneman, 2007).

## 2.5 ÍTEMS IDIOSINCRÁSICOS

Esta metodología se basa en el análisis de las respuestas individuales a los ítems para detectar el faking. Concretamente, se trata de identificar los patrones de respuesta de los sujetos en condiciones faking y condiciones honestas ante un conjunto particular de ítems, catalogados como idiosincrásicos. Esta estrategia parte del planteamiento de que se pueden identificar ítems en adjetivos que, por su carácter ambiguo, dan lugar a un patrón atípico de respuesta en condiciones de faking. Al conocer cómo es el patrón de respuesta que siguen los evaluados cuando distorsionan estos ítems, se pueden detectar aquellas personas que están cometiendo faking.

Kuncel y Borneman (2007) pusieron a prueba esta metodología, y crearon una escala de 11 adjetivos idiosincrásicos a partir de la escala de Goldberg (1992; The Goldberg Adjective Markers). Al aplicar esta escala en un estudio experimental encontraron que es un método efectivo para detectar el faking, concretamente detectan entre un 20% y un 37% de respuestas falsificadas con un porcentaje de error del 1% (falsos positivos). Asimismo, se trataría de un método resistente al entrenamiento ya que al no centrarse en los patrones de puntuaciones extremas es difícil que los sujetos sean instruidos para otro tipo de patrones más ambiguos. No obstante, a pesar de los resultados favorables, los autores también plantean ciertas limitaciones en el procedimiento. Primero, dado que se trata de una primera aproximación a la aplicación de este tipo de estrategia existe la posibilidad de que los resultados obtenidos se hubiesen visto afectados por las características de la medida utilizada, la extensión y la composición de los ítems del test podrían haber hecho más efectiva la medida que otros formatos. Segundo, también es importante tomar estos resultados con cautela ya que no se ha estudiado si el uso de esta herramienta afecta a la validez predictiva de la medida, ya que la mayoría de los métodos que se aplican para detectar el faking de modo directo atenúan la validez de criterio. Por todo lo dicho, estos autores indican la necesidad de testarlo en otras muestras, trabajos y con otros ítems (Kuncel y Borneman, 2007).

## 2.6 DETECCIÓN A TRAVÉS DEL LENGUAJE

Dentro de esta categoría podemos encontrar diferentes técnicas que se fundamentan en el análisis de la escritura de los candidatos. La frecuencia en el uso de categorías particulares de palabras muestra la relevancia de ese conjunto de palabras para

el individuo, revelando información psicológica de los sujetos como son las características personales (Ventura, 2012).

Varios autores han considerado que esta metodología se podría aplicar para detectar si han cometido faking. Analizando la posible relación entre el engaño y el lenguaje, Pennebaker y Francis (1996) encontraron que las personas que responden honestamente tienen unos patrones de escritura diferentes a los obtenidos en personas que estaban distorsionando sus respuestas, es decir, que el faking en la comunicación escrita podría ser detectado a través de análisis lingüísticos.

Aunque estas técnicas se han diseñado para analizar el faking en otras áreas, se han propuesto diferentes adaptaciones para su aplicación en la detección del faking en la evaluación de la personalidad. Una de las principales ideas es adaptar la estructura de los cuestionarios de personalidad o de cuestionarios biográficos y transformarlos en preguntas abiertas. Por ejemplo, para evaluar el factor conciencia se podría presentar una pregunta como “*Describe un día típico en su trabajo*”. Este planteamiento obliga a los sujetos a describir una “historia” que permitirá analizar sus patrones de escritura y detectar el faking por el lenguaje que han utilizado.

Podemos destacar tres técnicas que se podrían aplicar en la detección del faking siguiendo este planteamiento:

A. *Linguistic Inquiry Word Count* (LIWC; Pennebaker, Francis y Booth, 2001). Este software de análisis del lenguaje examina la proporción de palabras que se usan en el texto de las categorías que se han prefijado como de interés. En el caso del faking se diferenciaron tres categorías que reflejan patrones de falsificación (Newman, Pennebaker, Berry y Richards, 2003): (1) no hacer referencias en primera persona, como un mecanismo de distanciamiento del engaño que se está cometiendo, (2) el uso de palabras con connotaciones más negativas, como un reflejo de la culpabilidad que pueden sentir por estar engañando y (3) textos menos complejos, al centrar el esfuerzo cognitivo en crear la mentira los textos son menos elaborados morfológica y sintácticamente. Aquellas personas que estén falsificando sus escritos tendrán mayor ratio en estas tres categorías. En la práctica profesional se ha comprobado que este mecanismo es capaz de diferenciar los patrones de engaño en la comunicación.

B. *Latent Semantic Analysis* (LSA; Landauer y Dumais, 1997). Técnica muy similar a la anterior, aunque en este caso no se detectan categorías, sino que basa en crear relaciones según el significado semántico de las palabras. De tal modo que se pueden



generar vectores que representan constructos particulares, en nuestro caso, de la personalidad. Por ejemplo, en el texto se pueden detectar palabras como sociable, hablador, dominante, etc., que representarían el constructo extraversión. Aplicando esta metodología a la detección del faking, la técnica LSA puede clasificar el lenguaje relacionado con el faking creando un conjunto de vectores que representarían categorías semánticas indicativas de faking. Si los sujetos están cometiendo faking los vectores que generen en su comunicación correlacionan con los vectores que representan esta conducta.

C. *N-Grams*. A diferencia de las anteriores técnicas que analizan individualmente las palabras, en este caso propone examinar combinaciones de palabras para detectar si son composiciones más asociadas al faking (Ventura, 2012).

No obstante, esta metodología también presenta ciertas limitaciones. Se trata de una técnica que apenas se ha aplicado en el contexto de la SP, por lo que se desconoce su efectividad real contra los efectos del faking. Asimismo, es necesario realizar adaptaciones de los cuestionarios a estos programas y requiere tener a disposición recursos electrónicos suficientes para poder hacer uso de estos.

## 2.7 TÉCNICAS TRI

Esta metodología basada en la Teoría de la Respuesta al Ítem (TRI) es una técnica estadística que detecta el faking a través del análisis de patrones de respuesta. Se fundamentan en el análisis de las diferencias en los procesos de respuesta a los ítems comparando las respuestas entre muestras honestas y faking. Si se identifica el patrón de respuesta en condiciones de faking, a través de la aplicación de un conjunto de modelos matemáticos, se pueden obtener los parámetros entre los que variarían las respuestas en esta condición (Kuncel y Borneman, 2007; Zickar y Drasgow, 1996). Aunque este método ha tenido cierta repercusión (Stark et al., 2001), lo cierto es que se ha mostrado inconsistente y menos efectivo que otros métodos aquí descritos (Kuncel y Borneman, 2007).

## 2.8 MÉTODOS DE CORRECCIÓN DEL FAKING

Todas las medidas descritas para detectar el faking han sido ampliamente aplicadas, no obstante, se trata de estrategias que sólo permiten la identificación de *fakers*,



por lo que su uso se presume ligado a otras prácticas para evitar que los sujetos que han distorsionado sus respuestas perjudiquen la toma de decisiones. Algunos trabajos han revisado este tipo de medidas conocidas como métodos de corrección del faking (Reeder y Ryan, 2012), entre los que se pueden distinguir cuatro tipos:

1. *Métodos de ajuste de las puntuaciones*. Se trata de la modificación de las puntuaciones en base a un parámetro estipulado, este puede ser un ajuste subjetivo u objetivo a través de la aplicación de una fórmula matemática, por ejemplo, una ecuación de regresión.

2. *Eliminación de casos*. Consiste en excluir del proceso los sujetos que han sido clasificados como *fakers* con los métodos de detección del faking

3. *Pruebas de verificación o reevaluación*. Los sujetos catalogados como *fakers* tendrían que volver a cubrir los test.

4. *Interpretación cautelosa de los casos*. No se eliminan ni modifican las puntuaciones, sino que se analizan con precaución para determinar si son candidatos igualmente idóneos o no siguen en el proceso de selección.

La aplicación de algunas de estas técnicas correctivas no está exenta de polémica. Como hemos indicado la principal crítica que se puede asociar a todas las estrategias descritas se refiere a los criterios que aplican para discernir quién está cometiendo faking y quien no. En muchos casos se podría llegar a catalogar como *fakers* a candidatos que simplemente por su naturaleza puntúan alto en factores que se asocian a un perfil “típico” de faking dando lugar a falsos positivos. Por otra parte, el hecho de optar por ajustar las puntuaciones tampoco supone una solución viable ya que el ajuste artificial de los resultados puede, igualmente, no corresponderse con las verdaderas puntuaciones de los sujetos lo que sólo tendría como consecuencia una alteración de la medida, de su validez y fiabilidad (Salgado, 2005; Ones y Viswesvaran, 1998; Táuriz, 2011).

Además, para muchos autores el hecho de que los participantes obtengan puntuaciones altas en medidas que detectan el faking no es motivo suficiente para eliminarlos del proceso de selección (Reeder y Ryan, 2012). Se fundamentan en la idea de que el hecho de que cometan faking no implica que sean sujetos inadecuados para el puesto de trabajo. Es más, algunos posicionamientos defienden que puede que sean mejores candidatos ya que muchos investigadores han planteado que la capacidad de cometer faking está asociada con una mayor capacidad mental general de los sujetos. De

tal modo que las personas más inteligentes también son las que más fácilmente modificarían sus puntuaciones (Vasilopoulos, Cucina, Dyomina, Morewit, y Reilly, 2009) y la literatura ha demostrado la habilidad mental general es el mejor predictor del futuro desempeño del trabajador (Ones, Viwesvaran y Dilchert, 2005; Schmidt, 2002; Salgado, 2017). Sin embargo, esta relación ha sido poco estudiada.

### 3. MÉTODOS DE CONTROL DEL FAKING

Los métodos descritos hasta el momento permiten únicamente detectar el faking una vez este se ha cometido. Lo que requiere la aplicación de otras medidas correctivas para descartar a los sujetos que han falseado la aprueba y evitar que afecte a la toma de decisiones (Reeder y Ryan, 2012). No obstante, se ha desarrollado otro conjunto de métodos considerados como estrategias preventivas. Son medidas cuyo diseño permite reducir el grado en que los candidatos falsifican las pruebas, es decir, son estrategias que evitan que los sujetos caigan en conductas de faking.

Dentro de esta categoría se encuentran medidas como las estrategias de advertencia, los test de razonamiento condicionado o los cuestionarios de EF. Estos últimos no se describirán en este apartado, se ha preferido dedicarles un capítulo específico al ser el estudio de su efectividad contra el faking uno de los objetivos centrales de esta tesis doctoral.

#### 3.1 ESTRATEGIAS DE ADVERTENCIA

Es una de las medidas de reducción del faking más sencillas, pero a la vez se ha considerado una de las más eficaces (Dwight y Donovan, 2003). Se trata de informar a los candidatos, antes de que comiencen las pruebas, sobre las posibles penalizaciones que conllevaría si se detectase que han distorsionado sus respuestas (Salgado, 2005; Táuriz; 2011). Podemos diferenciar varios tipos de advertencia dependiendo de la información que se les facilite a los participantes.

Dwight y Donovan (2003) distinguieron tres tipos: (a) *advertencia de identificación*, cuando se informa de que se puede detectar a aquellas personas que distorsionan durante el proceso; (b) *advertencia sobre las consecuencias* del faking, cuando se exponen las acciones disciplinarias que se llevarían cuando alguien comete faking o (c) una combinación de ambas. Por su parte, Dilchert y Ones (2012)

distinguieron hasta cinco tipos de advertencia: (a) *advertencia de detección* y (b) *advertencia sobre las consecuencias*, equivalentes a la propuesta de Dwight y Donovan (2003); (c) *apelación a la razón*, en este caso se busca alterar la concepción del faking como algo positivo y que los sujetos entiendan que decir la verdad supone consecuencias más positivas para ellos; (d) *advertencia educativa*, se trata de explicar la naturaleza del proceso y que la precisión en las respuestas es beneficioso para todas las partes, empresa y candidatos y (e) *apelación a la moral*, se centra en las convicciones morales de los sujetos para explicar que la honestidad en las respuestas es lo moralmente correcto.

Las estrategias de advertencia son una de las medidas consideradas como más efectivas para la reducción del faking, altamente eficaz y sin efectos negativos hacia las percepciones de los solicitantes sobre la organización (Salgado, 2005). Varios estudios han demostrado que esta estrategia logra una reducción significativa del faking. McFarland (2003), Robson, Jones y Abraham (2008) y Alonso y Táuriz (2010) constataron que se produce una reducción ( $d$ ) promedio de .45, .35 y .60, respectivamente. Landers, Sackett y Tuzinski (2011) encontraron que la advertencia reduce aproximadamente un 60% la intención de cometer faking en un contexto real de evaluación. Sin embargo, algunos autores recalcan que su efectividad puede estar asociada a que la mayoría de estos estudios se han realizado en contextos experimentales (Dilchert y Ones, 2012).

Dwight y Donovan (2003) realizaron un meta-análisis para investigar el verdadero impacto de esta medida sobre el faking. En una primera comparación de los tamaños del efecto ( $d$ ) encontraron diferencias sustanciales entre los estudios incluidos en el meta-análisis, lo que reflejaba una inconsistencia en la efectividad de esta medida. No obstante, se obtuvo una  $d$  promedio de .23. lo que sugería que, en términos generales, la advertencia puede reducir el faking. Sin embargo, observaron que los resultados en los estudios primarios podrían estar condicionando los resultados del meta-análisis ya que sólo 5 de los 11 estudios primarios reportan un tamaño del efecto positivo. Por otro lado, Dwight y Donovan (2003) analizaron si el tipo de advertencia podría influir sobre la eficacia de esta medida. Distinguieron dos tipos de advertencia: (a) advertencia que informaba de que el test utilizado detectaba el faking y (b) advertencia sobre las consecuencias de distorsionar las respuestas. Los resultados mostraron que la segunda técnica ( $d = 0.30$ ) o una combinación de ambas ( $d = 0.25$ ) eran los métodos más efectivos. Por último, estos autores llevaron a cabo un estudio primario sobre la efectividad de la advertencia para reducir el faking. Los resultados apoyaron los obtenidos en su revisión. La estrategia de

advertencia reduce la conducta de faking y, aunque este efecto se puede considerar pequeño, ha resultado más efectivo que otros métodos basados en EDS e *ítems bogus*. Asimismo, el tipo de advertencia afecta sobre la efectividad de la medida, no obstante, tomados en conjunto, los tres tipos analizados reducen la presencia de faking.

Si esta técnica reduce del faking, debería producirse un aumento de la validez de las medidas no cognitivas. Sin embargo, los efectos de esta estrategia sobre la validez son menos convincentes. Robson et al. (2008) no encontraron un aumento de la validez convergente al utilizar esta estrategia, mientras que McFarland (2003) afirmó que la validez de las medidas aumentaba tras observar que con el uso de esta estrategia las correlaciones entre las dimensiones se reducían, es decir, que cada dimensión realizaba una mayor aportación singular en la predicción del constructo deseado, en este caso el desempeño.

Táuriz (2011) también llevo a cabo una investigación en la que analizó los efectos de esta estrategia sobre la validez predictiva de la personalidad. Comparó los resultados en tres condiciones, honesta, faking y de advertencia. Encontró que la condición de advertencia no evitó los efectos de atenuación que provoca el faking sobre la validez de la medida. En ese sentido, concluye que esta estrategia no permite recuperar la validez perdida como consecuencia del faking.

### 3.2 TEST DE RAZONAMIENTO CONDICIONADO

James (1998) presentó los problemas de razonamiento condicionado como una medida de control de los sesgos de autopresentación. Este tipo de pruebas se caracteriza por presentar problemas que requieren inducir inferencias a partir de la información ofrecida para seleccionar la respuesta correcta. Cada problema de razonamiento condicional propone dos respuestas inducibles a partir de la información, la elección de los sujetos indicará si están aplicando un mecanismo de justificación o están basando su respuesta en algún tipo de sesgo. Lebreton et al. (2007) propusieron este tipo de medidas como una herramienta adecuada para el control del faking; al obligar a los sujetos a centrar su atención en inducir la respuesta correcta provocan que la atención para responder de modo socialmente deseable decrezca. Estos mismos autores analizaron en la práctica la eficacia de esta medida y encontraron que, en condiciones en las que no se les explica la naturaleza real de la prueba, es decir, no se indica que es una medida indirecta de los sesgos, la prueba no es susceptible al faking.

Si bien se trata de una herramienta que aún está en desarrollo, estos resultados alientan a los investigadores de la SP a seguir explorando el uso de medidas indirectas para reducir el faking.

En lo que respecta a este capítulo se puede concluir que:

1. Se ha desarrollado un amplio conjunto de estrategias con el fin de detectar los efectos perjudiciales del faking que han alcanzado diferentes niveles de respaldo y aplicabilidad. Estas metodologías se pueden dividir en dos grandes grupos:

a) Aquellas que únicamente permiten discernir cuándo un sujeto está cometiendo faking, conocidas como estrategias de detección del faking. En esta categoría se incluyen las EDS, el análisis del tiempo de latencia o el uso de baremos específicos para muestras de candidatos.

b) Las estrategias conocidas como de reducción del faking, su característica principal es que están ideadas para mitigar las respuestas distorsionadas, de tal modo que se reduce la capacidad o la intención de los sujetos de cometer faking. La advertencia o los test de razonamiento condicionado se incluyen dentro de esta categoría.

2. De los métodos descritos, la advertencia previa a los candidatos, así como los baremos basados en muestras de candidatos parecen ser los mejores métodos para controlar los efectos del faking (Salgado, 2016). No obstante, los cuestionarios de EF, que detallaremos en el siguiente capítulo, han surgido como un método para superar las críticas hacia los aquí descritos.

## CAPITULO 5

# CUESTIONARIOS DE ELECCIÓN FORZOSA Y EL MODELO PSICOMÉTRICO DE LOS EFECTOS DEL FAKING





En el capítulo precedente se ha enumerado una importante variedad de métodos y estrategias que se han aplicado con el fin de disminuir la ocurrencia de este fenómeno, pero que, sin embargo, no han alcanzado el éxito deseado debido a las importantes limitaciones que presentan. En ese sentido, no se ha cesado en la búsqueda de mecanismos para reducir el impacto del faking en los procesos de evaluación y, especialmente, sobre las medidas de personalidad, siendo los cuestionarios de personalidad de EF una de las herramientas sobre las que más se ha investigado en los últimos años.

En este capítulo se realiza una revisión sobre los cuestionarios de personalidad de EF y su utilidad como medida de reducción del faking con el objetivo de comprender en profundidad el mecanismo que subyace a esta estrategia y su importancia como medida contra el faking.

## 1. LOS CUESTIONARIOS DE ELECCIÓN FORZOSA

Los cuestionarios de EF se han posicionado como una estrategia de reducción del faking en las medidas de personalidad, tratando de superar las críticas asociadas a otras medidas de control de este fenómeno. Su aparición se remonta a los años cuarenta (ver Gordon, 1951; Hicks, 1970; Zavala, 1965). Sin embargo, es ahora cuando ha despertado más el interés de la SP. Tett, Christiansen, Robie y Simonet (2011) encontraron que el 30% de las organizaciones utilizaban cuestionario de EF, pasando de ser una estrategia prácticamente olvidada a convertirse, en las últimas décadas, en una herramienta de referencia para combatir los efectos del faking.

### 1.1 CARACTERÍSTICAS DE LOS CUESTIONARIOS DE EF

El formato de EF surgió como medida para evaluar la personalidad hace más de setenta años. La idea original sobre este método fue concebida por Horst, mientras desarrollaba escalas de personalidad y por Wherry, quién planteó una idea similar aplicando medidas de personalidad en personal aeronáutico. No obstante, este primer acercamiento no se desarrolló en profundidad hasta los años cincuenta cuando se creó el primer test de personalidad estandarizado de EF, el *Edwards Personal Preference Schedule* (EPPS; Edwards, 1954), al mismo tiempo que proliferó la publicación de trabajos que realizaban un análisis crítico de este nuevo formato de medida de personalidad (ver Travers, 1951; Zavala, 1965).

La principal característica de los cuestionarios de EF es que cada ítem agrupa varias alternativas de respuesta, generalmente pares, tríadas o tétradas, de similar grado de deseabilidad, de tal modo que es más compleja la tarea de seleccionar la respuesta más aceptable socialmente. Se basan en el supuesto de que debido a la dificultad que tendrá el candidato para responder, éste siempre elegirá aquella alternativa que más lo identifique, reduciéndose así el efecto del faking sobre las puntuaciones (Christiansen, et al., 2005; Converse et al., 2006; Jackson, et al., 2000). Por lo tanto, los cuestionarios de EF se distinguen de las medidas SS (p. ej., escalas tipo Likert, verdadero/falso, grado de acuerdo, sí/no, etc.) en que los sujetos tienen que hacer una elección “forzada” entre varios ítems y no sólo calificar independientemente cada alternativa como ocurre en las medidas SS (Salgado, 2017).

Generalmente, las instrucciones de este tipo de cuestionarios obligan a los candidatos a escoger, entre las alternativas de cada ítem, aquella opción que “*más le defina*” o con lo que más identificado se sienta y/o la opción que “*menos le defina*” o que menos refleje su manera de ser. Un ejemplo de ítem de un cuestionario de EF sería:

*Soy una persona...*

- a) con grandes aptitudes*
- b) a la que le encantan los deportes acuáticos*
- c) que aprovecha todas las oportunidades para aprender*

Como se puede comprobar en el ejemplo, todas las alternativas son altamente deseables por lo que supondrá un mayor esfuerzo para el candidato deducir cuál de las opciones es la más relevante para el propósito de la evaluación, consecuentemente, como se presume, los sujetos tenderán a reflejar su propia personalidad en las respuestas. No obstante, como ocurre con otras estrategias de evaluación de la personalidad, se pueden distinguir diferentes formatos de cuestionarios de EF en función del diseño y procedimientos metodológicos aplicados para obtener las puntuaciones en cada una de las categorías.

### *1.1.1 Tipos de ítems*

Dentro de los cuestionarios de EF se diferencian dos tipos de ítems: *multidimensionales* y *unidimensionales*.

A. Los ítems *multidimensionales*. Son aquellos en los que cada una de las alternativas que componen cada pregunta refleja una dimensión o factor de la

personalidad. Las medidas ipsativas y quasi-ipsativas están compuestas por este tipo de ítems.

B. Los ítems *unidimensionales*. Están formados por conjuntos de afirmaciones que reflejan diferentes niveles de una única dimensión. Los ítems unidimensionales se asocian exclusivamente a medidas normativas. Algunos autores han puesto de manifiesto la dificultad de desarrollar medidas unidimensionales ya que resulta muy complejo reflejar diferentes grados de un constructo, especialmente cuando se trata de personalidad, siendo una tarea muy crítica y subjetiva. Asimismo, al tratarse de ítems que hacen referencia al mismo rasgo, podría ser más sencillo para los sujetos deducir qué está midiendo el ítem y, por consiguiente, reducir la eficacia de esta medida contra el faking (Converse et al., 2006).

### 1.1.2. Tipos de medida

Cuando nos referimos a los cuestionarios de EF debemos tener en cuenta que también se pueden distinguir tres categorías en función de las propiedades metodológicas que posean, es decir, de cómo se obtengan las puntuaciones para los factores que mide la prueba (Bartram, 1996; Clemans, 1966; Hicks, 1970). Concretamente, se pueden diferenciar medidas normativas, ipsativas y quasi-ipsativas o también denominadas parcialmente ipsativas.

A. *Medidas normativas*. Las medidas de EF normativas permiten realizar comparaciones a nivel de sujeto y entre grupos en cada una de las variables de personalidad que mide el test. Se trata, por tanto, de puntuaciones inter-individuales, de tal modo que los resultados que una persona obtenga en un factor evaluado por el test dependerá de los obtenidos en la población, pero serán independientes de las puntuaciones que obtenga el propio sujeto en otros factores. La característica que define a este tipo de cuestionarios es que cada ítem supone una elección en una escala bipolar que representa un único factor, es decir son ítems unidimensionales. Un ejemplo de test de personalidad de EF normativo sería el Myers-Briggs Type Indicator (MBTI; Myers, McCaulley, Quenk y Hammer, 1998).

B. *Medidas ipsativas*. Las medidas ipsativas son el segundo tipo de formato de EF. Clemans (1966) aclaró que una puntuación es ipsativa “cuando la suma de los

*atributos medidos para cada sujeto es constante*” (pág. 4). Los cuestionarios ipsativos se caracterizan por que todas las alternativas de cada ítem reciben una puntuación. En cada bloque el sujeto debe puntuar cada alternativa comparándola con las demás alternativas que presenta el ítem. En otras palabras, la puntuación que un sujeto obtiene en un factor depende de la puntuación que obtenga en los restantes factores que mida el test. Por lo tanto, se trata de puntuaciones intra-individuales, los resultados son dependientes a nivel individual pero independientes de las puntuaciones que otros sujetos obtengan. Este formato únicamente permite analizar los resultados a nivel de sujeto, pero se debe tener en cuenta que sólo nos muestra la importancia relativa de cada factor, no la puntuación real que posee un individuo. Es precisamente por este motivo por el que no se recomienda el uso de este tipo de medidas en contextos en los que sea necesario realizar una comparativa o ranking de todos los participantes, como es el caso de los procesos de selección, debido a que la información que partiríamos con esta medida estaría sesgada (Heggstad, Morrison, Reeve y McCloy, 2006). En esta categoría podemos encontrar varios test de personalidad ampliamente conocidos en la práctica profesional de la SP como, por ejemplo, el *Occupational Personality Questionnaire* (OPQ; SHL, 1993), *Edwards Personal Preferences Schedule* (EPPS; Edwards, 1957) o el *Description en Cinq Dimensions* (D5D; Rolland y Mogenet, 2001).

C. *Medidas quasi-ipsativas*. Los cuestionarios de EF quasi-ipsativos suponen una tipología que incluye medidas que no cumple con todos los criterios que se asocian a las medidas ipsativas, aunque comparten características con éstas. Hicks (1970), al comparar las propiedades de todos los formatos de EF, detalló que para que una medida sea quasi-ipsativa debe cumplir alguno de los siguientes requisitos en el diseño de la prueba: (1) los sujetos no tienen que ordenar todas las alternativas de cada ítem, lo hacen sólo parcialmente; (2) las escalas tienen diferente número de ítems; (3) no todas las alternativas clasificadas por lo candidatos son puntuadas; (4) las escalas se califican de forma diferente para los sujetos utilizando distintas transformaciones normativas en función de sus características; (5) las alternativas seleccionadas se ponderan de modo diferente; (6) algunas escalas ipsativas pueden ser eliminadas cuando se analizan los datos y, por último, (7) los tests tienen secciones normativas.

Hicks (1970) y Meade (2004) indican que las puntuaciones quasi-ipsativas se definen por las siguientes condiciones: (a) los resultados para cada factor varían entre los individuos a lo largo de un cierto rango de puntuaciones; (b) las puntuaciones no suman la misma constante para todas las personas, a pesar de que los test posean propiedades en

común con las pruebas ipsativas y (c) el aumento de la puntuación en un factor no necesariamente produce una disminución de la puntuación en otros factores. Pruebas como el *Gordon Personal Profile-Inventory* (GPPI-I; Gordon y Chevrier, 1993), el IPIP-MFC (Heggestad, Morrison et al., 2006) o el más reciente QI5F-tri de Salgado (2014) son ejemplos de cuestionarios quasi-ipsativos.

En resumen, los cuestionarios quasi-ipsativos comparten propiedades con las medidas normativas e ipsativas puras. Esta singularidad hace que las puntuaciones obtenidas con este formato de EF se puedan analizar a nivel intra e inter-individual (Salgado, 2017). En otras palabras, los resultados permiten conocer las diferencias individuales de cada sujeto y al mismo tiempo nos aportan información sobre las diferencias respecto a un grupo de referencia, lo que las hace que estas medidas sean más adecuadas que las ipsativas para la investigación desde el punto de vista estadístico.

No obstante, no debemos obviar que al tratarse de una medida *quasi*, se mantiene cierto grado de dependencia entre las diferentes escalas (factores) que componen este tipo de cuestionarios (McCloy Heggestad y Reeve, 2005) y es esta dependencia entre los factores lo que ha generado el mayor número de críticas asociadas a este tipo de medidas. Se han encontrado algunas desventajas metodológicas que afectan especialmente a las propiedades psicométricas de los cuestionarios de EF y que consideramos necesario abordar en el siguiente apartado (Carvajal y Gómez, 2014; Clemans, 1966; Converse et al., 2006; Dilchert y Ones, 2012; Hicks, 1970; Zavala, 1965, entre otros).

## 1.2 PROPIEDADES PSICOMÉTRICAS DE LOS CUESTIONARIOS DE EF

Las propiedades psicométricas han sido una de las mayores críticas que se han asociado a las medidas de EF. Autores como Hicks (1970), Baron (1996) o Bartram (1996) han publicado importantes trabajos analizando las propiedades psicométricas de este tipo de medidas y, especialmente, comparando las características de las medidas ipsativas con las medidas normativas. En términos generales las conclusiones a las que llegan estos autores son semejantes y ponen en entredicho la utilidad de los cuestionarios de EF como una medida para evaluar la personalidad.

La principal limitación asociada a las medidas ipsativas se deriva de la interdependencia de las escalas que la conforman. Como hemos explicado, los cuestionarios ipsativos se caracterizan por tener una restricción en las puntuaciones de suma constante, es decir, el conjunto de respuestas del test siempre sumará el mismo valor

(Meade, 2004). Esto implica que las puntuaciones entre las diferentes escalas (factores), necesariamente, se compensan entre sí, de tal modo que, si un sujeto puntúa alto en un factor, puntuará bajo en los restantes. Esta particularidad de los datos ipsativos da lugar a algunas propiedades psicométricas no deseables que, en cambio, no afectan a las medidas normativas:

1. La más importante de las propiedades a las que afectan las puntuaciones ipsativas es la correlación entre los factores. Concretamente, el formato ipsativo provoca que las intercorrelaciones de los factores que mide el test tiendan a ser artificialmente negativas (Hicks, 1970; Meade, 2004; Converse et al., 2006; Dilchert y Ones, 2012). Hicks (1970) ejemplificó este efecto de los test de EF ipsativos: suponiendo un test normativo que mida dos factores, A y B, se asume que la correlación entre los factores será próxima o igual a 0.00, dado que las puntuaciones en cada factor son independientes entre sí. Sin embargo, en el formato ipsativo, los candidatos al verse forzados a escoger la opción que representa al factor A o la que representan al factor B produce que las puntuaciones entre los factores A y B tengan una correlación  $r_{ab} = -1.00$  en lugar de 0.00. Por lo tanto, los resultados ipsativos generan correlaciones negativas entre los factores. Aunque, a medida que aumentan los factores que mide el test el efecto es menos pronunciado, la precisión de los resultados se va a ver afectada.

Esta limitación, a su vez, está asociada a otros problemas desde el punto de vista psicométrico. Uno de los más relevantes es que los test ipsativos carecen de una matriz de covarianzas interpretable. Clemans (1966) demostró matemáticamente que las sumas de las columnas o filas de una matriz de covarianza ipsativa serán iguales a cero y que las sumas de las columnas o filas de una matriz de intercorrelación ipsativa serán iguales a cero si las desviaciones ipsativas son iguales.

En cambio, las puntuaciones quasi-ipsativas reducen estas limitaciones. En este tipo de cuestionarios el sumatorio de las puntuaciones no es una constante, sino que se mueve entre un rango más amplio de valores. Este sistema de puntuación permite que los sujetos alcancen puntuaciones altas o bajas en más de un factor. Si bien, no es posible obtener puntuaciones muy altas en todos los factores, como sí puede ocurrir con los tests normativos, esta es una limitación pequeña desde el punto de vista empírico dado que un porcentaje muy pequeño de la población (menos del 5%) podría llegar a obtener puntuaciones muy altas en los cinco factores de personalidad. Pero, por el contrario, esta “flexibilidad” en las puntuaciones totales de los factores permite que se puedan interpretar



los datos como normativos y generar comparaciones inter-individuales, fundamentales en la SP.

2. La estructura factorial de las medidas de personalidad también se ve afectada cuando se emplean cuestionarios de EF. Las críticas se centran especialmente en las medidas ipsativas. Clemans (1966) y Meade (2004) plantearon que la dependencia de las variables en este tipo de medidas genera estructuras factoriales erróneas, creando factores bipolares artificiales debido a la presencia de correlaciones negativas entre las variables y a que sus sumatorios son siempre cero. Sin embargo, los estudios de Cattell y Brennan (1994) y, más recientemente, de Lee, Lee y Stark (2018) encontraron que la estructura factorial no se ve afectada en los cuestionarios de EF y que son los cuestionarios quasi-ipsativos los que mejores índices de ajuste presentan.

3. Por último, el estudio de la fiabilidad de los cuestionarios de EF ha generado un amplio debate en la literatura reciente. Tenopyr (1988) fue uno de los primeros investigadores en analizar la consistencia interna de las medidas de EF. Concluyó que el hecho de que las puntuaciones en los factores sean dependientes entre sí podía afectar artificialmente la fiabilidad de los factores, una fiabilidad muy alta en un factor puede aumentar la fiabilidad de los demás factores. El autor aclara que este efecto también se puede dar en la dirección opuesta. Savelli y Willson (1991), Baron (1996) y Bartram (1996), sin embargo, señalaron que los cuestionarios ipsativos tienden a obtener fiabilidades más bajas que los normativos. Salgado y Táuriz (2014) y Salgado et al. (2015) en recientes meta-análisis sobre la validez de las medidas de EF, aportaron evidencia empírica sobre esta cuestión llegando a la conclusión de que los efectos sobre los coeficientes de fiabilidad de las medidas de EF eran mínimos y no suponían ninguna implicación práctica.

No obstante, otros autores (Carvajal y Gómez, 2014; Dilchert y Ones, 2012; Meade, 2004) critican el uso de estadísticos clásicos como el alpha de Cronbach o el coeficiente test-retest para el cálculo de la fiabilidad, al considerarlos como inapropiados para este tipo de medidas. Se fundamentan en que el formato de EF no respeta las reglas básicas de la TCM (ver capítulo 3), es decir, la independencia de los factores. Por lo tanto, los resultados que se obtengan con estos estadísticos pueden ser erróneos y dar lugar a falsas conclusiones. Sin embargo, no se ha llegado a un consenso acerca de qué estadístico sería el más adecuado para su medición.

Las limitaciones descritas han llevado a muchos autores a criticar el uso de estas medidas en el ámbito profesional de la SP (Jackson et al., 2000). Sin embargo, no se debe



obviar que los cuestionarios de EF han surgido con la principal finalidad de controlar los sesgos asociados a este tipo de medidas. Hicks (1970) consideró importante indicar que los cuestionarios ipsativos deben usarse sólo en situaciones donde se ha demostrado que: (a) existe una respuesta de sesgo significativa; (b) este sesgo reduce la validez y (c) un formato ipsativo disminuye exitosamente el sesgo y aumenta la validez en mayor medida que lo harían otras medidas no ipsativas.

Otros autores también han considerado que las críticas a estas medidas pueden ser exageradas. Los investigadores han continuado analizando la capacidad predictiva y, sobre todo, la capacidad de reducir los efectos del faking que tienen los cuestionarios de EF frente a otras medidas de personalidad. A continuación, se realiza una revisión de los principales estudios y conclusiones que se ha obtenido.

### 1.3 VALIDEZ PREDICTIVA DE LOS CUESTIONARIOS DE EF

La validez de las medidas de EF ha sido una cuestión de gran relevancia para los investigadores. La principal utilidad de las medidas que evalúan la personalidad en la SP es la predicción del futuro desempeño en el trabajo. Por lo tanto, conocer la capacidad de los cuestionarios de EF para predecir comportamientos y criterios organizacionales resulta imprescindible.

Ha sido en las últimas décadas cuando se han publicado varios meta-análisis que evidencian la validez predictiva de las medidas de EF. Bartram (2005, 2007) llevó a cabo dos revisiones meta-analíticas que analizaron la validez predictiva de varias medidas ipsativas derivadas del *Occupational Personal Questionnaire* (OPQ; SHL, 1993). Tras analizar los resultados obtenidos en ambas revisiones, concluyó que las medidas de personalidad ipsativas OPQ resultan adecuadas predictoras del desempeño en el trabajo. Además, en el meta-análisis más reciente, encontró que el formato de los ítems tiene un ligero efecto moderador sobre la validez de las medidas. Concretamente, los resultados mostraron valores de validez más altos para las medidas ipsativas ( $p = .32/.38$ ) que para la medida tradicional con la que se compararon, en este caso una escala tipo Likert ( $p = .19/.25$ ).

Salgado y Táuriz (2014) examinaron la validez de los cuestionarios de EF para predecir criterios organizacionales y académicos. En el nivel más general encontraron que conciencia es el único predictor que generaliza su validez en el contexto organizacional, obteniendo una correlación corregida de  $p = .24$ . Estos resultados siguen

la línea de los obtenidos por Barrick y Mount (1991) para las medidas SS. En el caso de los criterios académicos, los resultados también mostraron que conciencia vuelve a ser el predictor más válido y generalizable de la personalidad con una  $p$  de 0.19. Aunque en este caso, la correlación es ligeramente inferior a la encontrada en meta-análisis de medidas SS, los autores justifican esta diferencia en el posible efecto de moderadores sobre la validez.

En cualquier caso, estos resultados demuestran que las medidas de personalidad en formato de EF son un válido predictor del desempeño organizacional y académico equiparable al de las medidas SS. Además, Salgado y Táuriz (2014) examinaron si el tipo de puntuación (normativa, ipsativa o quasi-ipsativa) modera la validez predictiva de los cuestionarios de EF. De los resultados se pueden extraer dos importantes conclusiones, la primera es que el formato quasi-ipsativo mostró ser un predictor más válido del desempeño que el formato ipsativo y, segundo, que las medidas quasi-ipsativas de conciencia son el mejor predictor del desempeño laboral ( $K = 44$ ,  $N = 8,794$ ,  $\rho = .40$ ), valor que prácticamente duplica al obtenido con las medidas ipsativas ( $\rho = .26$ ) y normativas ( $\rho = .22$ ) y que está por encima de cualquier otra medida SS, normativa o ipsativa de EF.

Salgado et al. (2015), como una extensión del meta-análisis anterior, publicaron un meta-análisis que analizaba el efecto moderador que tendría las categorías profesionales, es decir, los diferentes tipos de trabajo, sobre la validez de los cuestionarios ipsativos y quasi-ipsativos. Autores como Barrick y Mount (1991) o Salgado (1997) ya habían sugerido el potencial efecto moderador que tendrían los tipos de trabajo sobre la validez predictiva de las medidas SS. Sin embargo, apenas se había estudiado en medidas de EF. Salgado et al. (2015) analizaron la validez predictiva de los formatos quasi-ipsativos e ipsativos para predecir el desempeño de nueve categorías de puestos. Los resultados mostraron, de nuevo, que las medidas quasi-ipsativas tienen una mayor validez predictiva en cuatro de los cinco factores de personalidad, con la excepción de extraversión, a lo largo de todos los puestos analizados. En base a estos resultados, los autores concluyeron que las medidas quasi-ipsativas de personalidad eran los mejores predictores del desempeño para todos los grupos profesionales.

Más recientemente, Lee et al. (2018) estudiaron la validez de criterio de diferentes medidas de EF. Los resultados demuestran que el formato de EF quasi-ipsativo multidimensional es el método más efectivo para predecir el desempeño de los futuros

trabajadores. Estos resultados contribuyen a la evidencia empírica obtenida por los meta-análisis descritos de Salgado y Táuriz (2014) y Salgado et al., (2015).

Por último, Salgado (2017) analizó si la complejidad del puesto de trabajo moderaba la capacidad predictiva de los cuestionarios de personalidad de EF quasi-ipsativos. Concluyó que el nivel de complejidad era un robusto moderador de la validez de ambos formatos de EF. Cuanto mayor sea la complejidad del trabajo más se reduce la validez predictiva de las medidas de todos los factores con la única excepción del factor apertura a la experiencia. Por lo que se debe tener en cuenta este moderador al analizar la validez predictiva de las medidas de personalidad.

Los resultados descritos sobre la validez predictiva de los cuestionarios de EF tienen importantes implicaciones prácticas: (a) esta medida de personalidad es un predictor válido del desempeño a nivel organizacional y académico y (b) la validez de la medida quasi-ipsativa de conciencia es equiparable o superior a la obtenida con medidas SS. Por lo que se puede afirmar que las medidas de EF son una adecuada alternativa a las medidas SS para la evaluación y toma de decisiones académica y laboral.

## **2. CUESTIONARIOS DE ELECCIÓN FORZOSA Y FAKING**

Los cuestionarios de EF han sido creados con la finalidad de reducir los efectos nocivos del faking sobre las medidas de personalidad (Dichert y Ones, 2012; Jackson et al., 2000; Vasilopoulos et al., 2006, entre otros). Sin embargo, la justificación teórica de que este tipo de diseño es menos susceptible al faking debe ser respaldada empíricamente para que la utilidad de esta herramienta se demuestre y, en ese sentido, los resultados de los estudios primarios publicados sobre esta cuestión han sido inconsistentes.

Algunos estudios sugieren que no existen diferencias significativas entre utilizar una medida de EF o una tradicional. Es el caso del estudio de Heggstad, Morrison et al. (2006), quienes examinaron la resistencia al faking de una medida de EF quasi-ipsativa en comparación a medidas tradicionales tipo Likert. Encontraron que a nivel individual las puntuaciones quasi-ipsativas estaban afectadas por el faking en igual medida que los tests tradicionales. Heggstad, Morrison et al., (2006) concluyeron, por tanto, que el formato de EF no era un método adecuado para el control del faking.

Por el contrario, otras investigaciones han sugerido que los cuestionarios de EF son menos susceptibles a los efectos de faking. Jackson et al. (2000) y Christiansen et al. (2005) mostraron en sus respectivas investigaciones que el faking se reduce

sustancialmente cuando se utiliza una medida de EF. Jackson et al. (2000) no encontró diferencias significativas en las correlaciones entre las condiciones de baja y alta motivación a cometer faking cuando se aplicó el formato de EF ( $r_a = .41$ ;  $r_b = .36$ ;  $z = .60$ , n. s.). Christiansen et al. (2005) por su parte, concluyeron que, aunque el faking no se elimina, la distorsión en las respuestas es sustancialmente menor en los cuestionarios de EF ( $d = 0.26$ ) que cuando se utilizan medidas SS ( $d = 0.96$ ), inclusive en contextos donde la motivación para distorsionar es alta. Bowen, Martin y Hunt (2002) encontraron diferencias más pequeñas en las respuestas de los participantes ( $N = 301$ ) entre la condición honesta y la condición faking cuando usaron tests ipsativos ( $d = 0.24$ ), mientras que para el formato normativo las diferencias fueron significativas ( $d = 0.41$ ). Apoyándose en estos resultados, los autores concluyeron que el formato de EF es menos susceptible al faking. Sin embargo, los resultados descritos reflejan conclusiones de estudios primarios realizados en contextos controlados de laboratorio y con muestras reducidas, por lo que sus conclusiones deben ser tomadas con precaución.

Nguyen y McDaniel (2000) fueron de los primeros en realizar un meta-análisis para analizar si los cuestionarios de EF son más resistentes al faking que las medidas tradicionales de estímulo único. Los autores estudiaron los efectos a nivel inter e intra-grupo. Los resultados se resumen en la Tabla 11. Los resultados mostraron que las puntuaciones en los cuestionarios de EF podían ser distorsionadas, con  $d$  promedio de 0.40 y 0.58 para los diseños inter e intra-grupo respectivamente. No obstante, estos valores promedio fueron menores que los reportados por Viswesvaran y Ones (1999) para medidas SS, lo que indica que las medidas de EF pueden reducir los efectos del faking. Sin embargo, este meta-análisis tiene muchas limitaciones metodológicas que pueden estar influyendo en los resultados obtenidos. La primera de ellas es que se trata de un meta-análisis *bare-bones* por lo que no se han realizado las correcciones pertinentes y los datos resultantes pueden estar afectados por errores artificiales que no se han eliminado. Además, no se han analizado posibles moderadores que pueden estar afectando a los datos como, por ejemplo, el tipo de muestra o el tipo de faking (se han podido incluir datos de faking positivo y negativos conjuntamente). Estas limitaciones ponen en cuestión las implicaciones teóricas y prácticas que se podrían derivar de este meta-análisis.

Tabla 11. Resultados del Meta-análisis de Nguyen y McDaniel (2000) sobre los Efectos del Faking en los Cuestionarios de EF según el Tipo de Diseño

Diseño	$K$	$N$	$d$	$SD_d$	IC <sub>d</sub> 95%	
					Inf.	Sup.
<i>Intra-grupo</i>						
EE	8	422	0.49	.40	-0.07	1.06
EX	16	737	0.43	.31	0.33	0.53
AP	14	428	0.74	.21	0.11	0.14
A	12	355	0.76	.79	-0.40	1.91
C	32	1,350	0.59	.48	-0.10	1.28
Total	82	3,292	0.58	.46	-0.04	1.20
<i>Inter-grupos</i>						
EE	7	2,037	0.45	.23	0.07	0.84
EX	15	5,323	0.25	.16	0.02	0.48
AP	9	3,119	0.40	.23	0.00	0.79
A	11	4,396	0.45	.19	0.15	0.76
C	19	6,884	0.46	.25	0.00	0.91
Total	61	21,759	0.40	.23	-0.01	0.80

*Nota.* EE = estabilidad emocional; EX = extraversión; AP = apertura a la experiencia; A = amigabilidad; C = conciencia; *K* = número de tamaños del efecto; *N* = muestra acumulada; *d* = tamaño del efecto ponderado por tamaño de la muestra; *SD<sub>d</sub>* = desviación estándar del tamaño del efecto; IC<sub>d</sub> 95% Inf. / Sup. = valores inferior y superior del intervalo de credibilidad del 95%. Tabla adaptada de “*Faking and forced-choice scales in applicant screening: a meta-analysis*” de N. T. Nguyen y M.A. McDaniel, 2000, Paper of 15<sup>th</sup> Annual Conference of the Society for Industrial and Organizational Psychology.

Adair (2014) llevo a cabo una revisión meta-analítica para estimar la eficacia de diversos métodos para el control del faking, entre ellos, los cuestionarios de EF. De los resultados obtenidos concluyó que esta herramienta reduce en promedio la presencia del faking en los contextos de selección, en comparación al uso de cuestionarios de personalidad tradicionales, con tamaños del efecto que oscilan entre  $d = 0.15$  para amigabilidad y  $d = 0.70$  para el factor conciencia. No obstante, matiza que los tamaños de efectos obtenidos están condicionados por el tipo de diseño aplicado ya que en su mayoría los estudios llevados a cabo se realizan en contextos experimentales con unas características particulares que pueden diferir de los estudios correlacionales. Los resultados que obtuvo demostraron que las diferencias en las puntuaciones entre la condición honesta y faking se reducen cuando se tiene en cuenta este moderador. Asimismo, Adair (2014) indica que el tipo de muestra, inter-grupos o intra-grupo también puede afectar a los datos. Sin embargo, este autor no tiene en cuenta otro tipo de

moderadores en sus análisis que también pueden estar afectando a los resultados obtenidos como son el formato de los cuestionarios (normativo, ipsativo o quasi-ipsativo), ni aporta evidencia empírica sobre la resistencia al faking de los cuestionarios de EF en contextos reales de selección.

Cao (2016) y Cao y Drasgow (2019) han publicado el que es en la actualidad el meta-análisis más reciente y completo de los efectos del faking sobre las medidas de personalidad en formato de EF (Tabla 12).

Tabla 12. Resultados del Meta-análisis General de Cao y Drasgow (2019) de los efectos del Faking en los Cuestionarios de EF

	<i>K</i>	<i>N</i>	<i>d</i>	<i>SD<sub>d</sub></i>	<i>IC<sub>d</sub></i> 95%	
					Inf.	Sup.
EE	32	124,994	0.00	.08	-0.02	0.03
EX	40	169,069	0.16	.22	0.09	0.22
AP	36	217,100	0.00	.10	-0.03	0.03
A	38	260,288	0.19	.35	0.08	0.31
C	53	262,748	0.23	.35	0.13	0.32
Total	74	267,586	0.06	.15	0.02	0.09

*Nota.* EE = estabilidad emocional; EX = extraversión; AP = apertura a la experiencia; A = amigabilidad; C = conciencia; *K* = número de tamaños del efecto; *N* = muestra acumulada; *d* = tamaño del efecto ponderado por tamaño de la muestra; *SD<sub>d</sub>* = desviación estándar del tamaño del efecto; *IC<sub>d</sub>* 95% Inf. / Sup. = valores inferior y superior del intervalo de credibilidad del 95%. Tabla adaptada de “Does forcing reduce faking? A meta-analytic review of forced-choice personality measures in high-stakes situations” de M. Cao y F Drasgow, 2019, Journal of Applied Psychology.

Los resultados del meta-análisis general pusieron de manifiesto la reducción de los efectos del faking cuando se utiliza una medida de personalidad de EF. El factor conciencia, como era esperable, sigue siendo el más afectado por la distorsión ( $d = 0.23$ ) seguido de los factores amigabilidad ( $d = 0.19$ ) y extraversión ( $d = 0.16$ ). No obstante, en todos los casos los tamaños del efecto fueron inferiores a los obtenidos con medidas SS (Birkeland et al., 2006; Viswesvaran y Ones, 1999). Por lo tanto, estos resultados demuestran la efectividad de las medidas de EF para reducir el faking.

Asimismo, estos autores también estudiaron el efecto de varias variables moderadoras sobre la efectividad de las medidas de EF. Primero analizaron si el tipo de cuestionario (pick/normativo o mole/quasi-ipsativo) y el diseño de las preguntas (unidimensional o multidimensional) afectaba a la reducción del faking. A diferencia de los estudios previos, encontraron que las medidas normativas eran más efectivas contra



el faking ( $K = 40$ ,  $N = 263,096$ ,  $d = 0.05$ ) que las quasi-ipsativas ( $K = 32$ ,  $N = 3,852$ ,  $d = 0.37$ ). Sin embargo, este resultado puede estar afectado por los tamaños muestrales. La muestra de las medidas normativas es muy superior, lo que podría estar condicionando los resultados obtenidos incluso cuando los propios autores extraen de los análisis los estudios más amplios. Lo mismo ocurriría con los resultados sobre el tipo de ítem, el meta-análisis mostró que las medidas multidimensionales son más efectivas contra el faking,  $d = 0.03$  frente a una  $d$  de 0.13. para las unidimensionales, pero al realizar los análisis sin las muestras más grandes estas diferencias se eliminan, por lo tanto, los tamaños muestrales están afectando también a estas conclusiones.

Respecto a las otras variables moderadoras analizadas, hay que destacar que, en cuanto al tipo de diseño utilizado, se encontró que los estudios experimentales (faking inducido) muestran en promedio tamaños del efecto más altos que los estudios de campo (conciencia obtuvo el valor más elevado,  $d = 0.70$  en la muestra total y  $d = 0.18$  eliminando las muestras más grandes). Estos resultados, van en la línea de los obtenidos por Birkeland et al. (2006), aunque en el caso de las medidas de EF, los valores fueron significativamente menores que los hallados en las medidas SS. No se hallaron diferencias significativas entre los diseños intra-grupo e inter-grupos.

Sin embargo, a pesar de las importantes contribuciones que presenta este meta-análisis también debemos indicar algunas limitaciones metodológicas. La primera y más relevante son las diferencias en los tamaños muestrales de los estudios que incluye. Los datos de los meta-análisis generales provienen prácticamente de un único estudio (ver Drasgow et al., 2012) con importantes críticas metodológicas; la principal es que las condiciones experimentales no se ajustan a las condiciones de faking, es decir, no se trataría de una muestra de solicitantes tal y como es definida por Cao y Drasgow (2019), por lo que está condicionando todos los resultados que se deriven de su inclusión. Asimismo, en muchos casos los tamaños del efecto incluidos son promedios de escalas que los autores no aclaran cómo han calculado. Por tanto, falta información sobre los criterios de inclusión que se han utilizado para la realización de este meta-análisis.

Las limitaciones de los meta-análisis analizados generan la necesidad de llevar a cabo un nuevo meta-análisis que incluya tanto los estudios empleados como estudios no incluidos en los mismos. Asimismo, los estudios y meta-análisis realizados hasta la fecha han examinado los efectos del faking sobre las medias y desviaciones de las puntuaciones de los factores. Sin embargo, según la teoría psicométrica del faking, este fenómeno no sólo puede afectar a estos dos parámetros. Salgado (2016) encontró que el faking,



evaluado en medidas SS, además de aumentar la media y disminuir la desviación típica también afectaría a la fiabilidad, la estructura factorial y la validez de criterio. No obstante, hasta donde llega nuestro conocimiento estos efectos no se han examinado en los cuestionarios de EF, por lo que sería necesario examinar esta cuestión empíricamente para estimar los efectos sobre estas medidas. Para finalizar, se puede concluir que:

1. Los cuestionarios de personalidad de EF son un método evaluativo concebido para reducir los efectos del faking. No obstante, han sido ampliamente cuestionados por sus propiedades psicométricas, especialmente asociadas a las puntuaciones ipsativas:

a) La dependencia de las puntuaciones de los factores ha sido la principal crítica realizada a estas medidas. Esta dependencia produciría correlaciones negativas entre los factores y afectaría a la estructura factorial de la medida (Clemans, 1966; Hicks, 1970; Meade, 2004).

b) Varios autores han considerado que la fiabilidad de estas medidas también se vería afectada. Si bien, no existe un consenso en cuanto a si aumentarían o disminuirían los valores de los coeficientes de fiabilidad (Baron, 1996; Bartram, 1996; Heggstad et al., 2006; Savelli y Willson, 1991; Tenopyr, 1988).

c) Por último, también se ha puesto en duda la propia efectividad de la medida para controlar el faking (Jackson et al., 2000).

2. A pesar de estas críticas, la evidencia empírica más reciente ha demostrado que los cuestionarios de EF son una excelente herramienta para la evaluación del desempeño en contextos académicos y organizacionales.

a) Los meta-análisis de Bartram (2005, 2007), Fisher et al. (2019) y Salgado y Táuriz (2014) mostraron que las medidas de personalidad en formato de EF son un predictor válido del desempeño organizacional y académico.

b) Además, Salgado y Táuriz (2014) llegaron a dos importantes conclusiones, el formato quasi-ipsativo mostró ser un predictor más válido del desempeño que el formato ipsativo y las medidas quasi-ipsativas de conciencia son el mejor predictor del desempeño laboral.

c) Por su parte, el meta-análisis de Salgado et al. (2015) analizó el efecto moderador que podrían tener los tipos de puestos evaluados con estas medidas. Sus hallazgos mostraron que las medidas quasi-ipsativas de personalidad eran los mejores predictores del desempeño para todos los grupos profesionales.

3. En cuanto a la evidencia de la efectividad de los cuestionarios de EF para reducir el faking, el número de meta-análisis es muy escaso. Si bien, la principal conclusión a la que han llegado es que se trata de una herramienta adecuada para disminuir los efectos perniciosos del faking en los procesos selectivos. No obstante, es necesario tener en cuenta ciertas consideraciones:

a) Nguyen y McDaniel (2000) no realizaron correcciones de los efectos de los errores artificiales. Además, no examinó el efecto moderador del tipo de cuestionario de EF.

b) Adair (2014) también concluyó en su meta-análisis que los cuestionarios de EF reducen la presencia del faking en los procesos de selección. Sin embargo, estos resultados pueden estar afectados por ejemplo por el tipo de diseño.

c) Cao y Drasgow (2019) realizaron un trabajo más elaborado al exhaustivo el efecto moderador de un amplio conjunto de variables criterio. Si bien, no realizaron correcciones por falta de fiabilidad ni restricción en el rango. Asimismo, los criterios de inclusión de algunas de los estudios, así como sus tamaños muestrales podrían estar condicionando los resultados obtenidos.

4. Debido a las limitaciones metodológicas señaladas resulta necesario realizar un nuevo meta-análisis para conocer de un modo más preciso los efectos del faking sobre los cuestionarios de EF.

5. Los estudios sobre los cuestionarios de EF se han centrado en examinar los efectos del faking sobre las distribuciones de las puntuaciones de estas medidas y compararlos con los obtenidos al emplear cuestionarios SS. Sin embargo, no se han analizado los efectos del faking sobre la fiabilidad, la estructura factorial o la validez de criterio, que según la teoría psicométrica de los efectos del faking también se verían afectados (Salgado, 2016).

Por todo ello, la parte empírica de esta tesis doctoral pretende ser una contribución empírica sobre los efectos del faking, al poner a prueba el modelo teórico del faking sobre las medidas de EF y estimar la resistencia de estas medidas a este fenómeno.

# PARTE EMPÍRICA

EVALUACIÓN EMPÍRICA DEL MODELO  
TEÓRICO EN CUESTIONARIOS DE ELECCIÓN  
FORZOSA



# ESTUDIO 1

META-ANÁLISIS DE LOS EFECTOS  
PSICOMÉTRICOS DEL FAKING SOBRE LOS  
CUESTIONARIOS DE PERSONALIDAD DE  
ELECCIÓN FORZOSA



## INTRODUCCIÓN

En los capítulos teóricos de esta tesis hemos visto como una parte de la investigación empírica sobre el faking ha tratado de demostrar, mediante un amplio conjunto de estudios primarios, que los cuestionarios de EF reducen los efectos de este fenómeno en las medidas de personalidad. Sin embargo, son escasas las revisiones meta-analíticas llevadas a cabo para analizar la resistencia al faking de estas medidas de personalidad.

Realizar una revisión cuantitativa de los estudios primarios sobre esta cuestión es una tarea necesaria para estudiar la efectividad de los cuestionarios de EF ya que el meta-análisis permite resumir la evidencia empírica y obtener conclusiones más precisas y robustas sobre la cuestión estudiada.

En este caso, los meta-análisis de Adair (2014), Cao (2016; Cao y Drasgow, 2019) y Nguyen y McDaniel (2000) han sido los únicos que han aportado resultados integrados sobre la resistencia al faking de los cuestionarios de EF (ver Capítulo 5). No obstante, también presentan una serie de limitaciones que se han tratado de superar a la hora de realizar este estudio. De modo más específico, los aspectos que pueden ser mejorados se pueden resumir en los siguientes puntos:

1. *Características de las muestras y codificación de los datos.* Respecto al meta-análisis de Nguyen y McDaniel (2000) el número de estudios independientes que integra es muy reducido, lo que podría condicionar los resultados obtenidos. Además, su meta-análisis no habría diferenciado entre estudios con conductas de faking *good* y faking *bad*, integrando datos de ambas condiciones en la realización de los cálculos. Asimismo, la duplicidad de muestras ha sido otra posible limitación asociada a este trabajo. En cuanto al trabajo de Cao (2016; Cao y Drasgow, 2019), a pesar de su reciente publicación, incluye estudios independientes cuyas condiciones experimentales no se ajustarían a los estándares de conductas de faking. Es el caso del estudio de Drasgow, et al. (2012), incluido en su meta-análisis, el cuál combinó instrucciones de faking con instrucciones de advertencia. Esto podría haber afectado a la verdadera intención de faking y, por tanto, afectar la condición experimental y los resultados que se obtuvieron en el estudio, pero también los resultados del meta-análisis, al tratarse de una muestra muy grande en relación con el total de estudios incluidos en la revisión meta-analítica. Por último, el meta-análisis de Cao (2016; Cao y Drasgow, 2019) también omite información



sobre cómo se han codificados los datos, ya que únicamente se reporta una *d* promedio en estudios primarios cuando se están evaluando cinco factores independientes de personalidad.

2. *Corrección de artefactos*. La presencia de artefactos (p. ej., error de muestreo, error de medida y restricción en el rango) puede afectar a los datos, de tal modo que se obtengan resultados y conclusiones erróneas (Hunter y Schmidt, 1987 a y b), por lo que resulta conveniente corregir sus efectos. Tanto Adair (2014) como Cao (2016; Cao y Drasgow, 2019) y Nguyen y McDaniel (2000) llevaron a cabo meta-análisis de tipo *bare-bones*, por lo que la única fuente de error que han tenido en cuenta fue el error de muestreo. En ningún caso se ha considerado realizar correcciones por falta de fiabilidad (error de medida) o restricción en el rango en los meta-análisis llevados a cabo hasta la fecha.

3. *Análisis de moderadores*. Adair (2014) estudió el efecto moderador del tipo de diseño sobre la resistencia al faking de los cuestionarios de EF, aunque sólo aportó resultados a nivel de muestras de laboratorio. Cao (2016; Cao y Drasgow, 2019) ha estudiado el efecto moderador de dos conjuntos de variables: el tipo de test y el diseño experimental. No obstante, hasta el momento, ninguno de los meta-análisis ha considerado llevar a cabo un análisis jerárquico de los moderadores que han sido considerados de modo independiente en la literatura científica sobre el faking. Tanto el diseño experimental como el formato del test pueden afectar a los tamaños del efecto, por lo que es necesario considerar su efecto en conjunto y no por separado.

Estas limitaciones justifican la necesidad de llevar a cabo un meta-análisis para superar las limitaciones de los estudios anteriores con el objetivo de ampliar y mejorar los trabajos de Adair (2014), Cao (2016; Cao y Drasgow, 2019) y Nguyen y McDaniel (2000). Para ello se ha realizado este meta-análisis aplicando los métodos de meta-análisis psicométrico con corrección de artefactos desarrollados por Hunter y Schmidt (1990, 2004), para que los resultados obtenidos sean lo más robustos posibles.

## HIPÓTESIS DE ESTUDIO

Los meta-análisis que se presentan en este primer estudio tienen dos objetivos principales: (a) comprobar si los inventarios de personalidad de EF son robustos contra los efectos del faking y (b) analizar posibles variables moderadoras que puedan estar afectando la capacidad de los cuestionarios de EF para controlar los efectos del faking. Se analizan dos conjuntos de moderadores: por un lado, el diseño de los estudios acumulados y, por otro, el formato del test aplicado.

En base a la literatura presentada en la primera parte de esta tesis se puede concluir que el objetivo con que fueron desarrollados los cuestionarios de EF es la reducción de los efectos del faking sobre las respuestas de personalidad. Las características metodológicas de estas medidas permitirían controlar en mayor medida la intención de los sujetos para distorsionar sus respuestas (ver capítulo 5). Las integraciones previas sobre esta cuestión han apoyado este planteamiento (Nguyen y McDaniel, 2000; Cao, 2016; Cao y Drasgow, 2019). Por tanto, la hipótesis planteada es:

*H1: Los cuestionarios de EF tienen resistencia al faking.*

Asimismo, dentro de los cuestionarios de EF podemos distinguir tres tipos de medidas diferentes, test ipsativos, quasi-ipsativos y normativos. En la literatura se ha planteado que las diferencias metodológicas entre cada uno de los formatos también pueden afectar a su capacidad para reducir el faking. En ese sentido, el meta-análisis de Cao (2016; Cao y Drasgow, 2019) ha ofrecido una primera respuesta a esta cuestión, hallando que los cuestionarios normativos son los que mayor resistencia ofrecen al faking. Sin embargo, estos resultados pueden estar sesgados por los datos que incluyen en sus análisis.

Basándonos en las características que definen a cada formato, los cuestionarios de tipo quasi-ipsativo parecen ofrecer una mayor capacidad para reducir los efectos del faking, ya que se trata de medidas en las que se requiere un mayor esfuerzo cognitivo para seleccionar la respuesta adecuada, lo que puede reducir la intención de falsificar las respuestas. De este modo planteamos en relación con la hipótesis primera, que:

*H2: Los cuestionarios de EF quasi-ipsativos, por sus características metodológicas, son los que presentan mayor resistencia al faking.*

Estudios previos han demostrado que el tipo de diseño utilizado afecta a los resultados. Viswesvaran y Ones (1999) encontraron diferencias significativas entre los diseños intra-grupo e inter-grupos cuando analizaron el efecto del faking en medidas tradicionales de personalidad. Birkeland et al. (2006) mostraron que el faking se producía en menor magnitud en contextos reales de selección que en estudios experimentales y Salgado (2016) también halló diferencias significativas en la incidencia del faking entre muestras de candidatos, trabajadores y diseños intra-grupo. Siendo en los estudios experimentales, especialmente, los diseños intra-grupo, seguido de los estudios con muestras de candidatos, en donde se producía más faking. Por lo tanto, en base a la literatura descrita, las hipótesis que se plantean son:

*H3: El diseño de los estudios actúa como moderador de los efectos del faking en las medidas de personalidad.*

*H3a: Se produce una mayor incidencia del faking en contextos experimentales que en contextos reales de selección.*

*H3b: Las muestras intra-grupo, en los diseños experimentales, y las muestras de candidatos, en los diseños correlacionales, se ven afectados en mayor medida por el faking.*

## MÉTODO

### 1. BÚSQUEDA DE ESTUDIOS

Se realizó una búsqueda exhaustiva con el objetivo de obtener el mayor número de estudios que analizaran la resistencia de los cuestionarios de EF al faking o que publicasen datos que permitiesen estudiar este efecto. Para alcanzar este objetivo se emplearon cinco estrategias:

1. Se llevó a cabo una búsqueda en las bases de datos electrónicas: *EBSCO Host*, *PsycInfo*, *Researchgate*, *ScienceDirect*, *Taylor y Francis* y *Wiley Online Library* y en los buscadores *Google* y *Google Scholar*. Utilizando las palabras clave *forced-choice*, *forced-choice questionnaire*, *forced-choice format* o *ranking format* para la variable cuestionarios de EF y *faking*, *impression management*, *response distortion*, *social desirability* o *score distortion* para hacer referencia a la variable faking. Por último, también se emplearon términos que hiciesen referencia a la personalidad para ser más específicos en la búsqueda, utilizando

palabras clave como *personality* y *Big Five*. También se aplicaron las variantes en español de estos términos en la realización de la búsqueda.

2. Se realizó una búsqueda artículo por artículo en los números publicados en los últimos treinta años, hasta junio de 2019, en las siguientes revistas científicas: *European Journal of Work and Organizational Psychology*, *Human Performance*, *Journal of Clinical Psychology*, *Journal of Applied Psychology*, *Journal of Educational Measurement*, *Journal of Work and Organizational Psychology*, *Personality and Individual Differences*, *Personnel Psychology*, *Psychological Reports* y *Journal of Social Psychology*.

3. Se revisó la sección de referencias de los artículos encontrados en las revistas anteriormente mencionadas.

4. Se revisaron las referencias de los meta-análisis previos sobre faking y cuestionarios de EF: Nguyen y McDaniel (2000), Cao (2016), Cao y Drasgow (2019), Salgado y Táuriz (2014). Se trató de incluir exactamente los mismos estudios de las revisiones indicadas, aunque esto no fue posible debido a las limitaciones de acceso a seis de los estudios analizados en estas integraciones.

5. Por último, se contactó con investigadores del área para la obtención de estudios o datos no publicados. Se pidió información a un total de diez autores, obteniéndose los datos solicitados en cuatro de los contactos realizados.

## 2. CRITERIOS DE INCLUSIÓN

Una vez identificados y recopilados los estudios se procedió a su revisión para comprobar si cumplían los siguientes requisitos:

1. Los estudios debían proporcionar un indicador del tamaño del efecto sobre la relación de interés, en este caso, las diferencias en faking con el uso de cuestionarios de EF. También se incluyeron en el meta-análisis aquellos estudios que, a pesar de no reportar un valor del tamaño del efecto, proporcionasen datos que permitiesen su cálculo.

2. En cuanto a la condición experimental, se excluyeron del meta-análisis aquellos estudios en los que las condiciones de las investigaciones no se ajustasen a las condiciones de respuesta faking o de respuesta honesta. Específicamente, se descartaron aquellas investigaciones que incorporaron otro tipo de estrategias unidas a las instrucciones de faking, por ejemplo, instrucciones específicas de

advertencia sobre el proceso, ya que podrían estar sesgando los resultados. También se excluyeron estudios cuyas condiciones experimentales pusiesen en duda la honestidad de la muestra a la hora de responder.

3. Se incluyeron en la revisión meta-analítica estudios que utilizarasen cuestionarios de personalidad de EF en cualquiera de sus tres formatos (ipsativo, cuasi-ipsativo o normativo).

4. Por último, en cuanto a las variables de personalidad, se incluyeron todos los estudios cuyas medidas estuviesen basadas en el modelo de los cinco grandes factores o que se pudiesen clasificar dentro de este modelo. En este segundo caso, se aplicaron dos estrategias de clasificación: (a) se usaron las taxonomías de Birkeland et al. (2006) y de Salgado y Táuriz (2014) y (b) se consultaron las definiciones de las variables de personalidad proporcionadas por los manuales de las pruebas para, mediante a la consulta con expertos, consensuar la clasificación de las medidas de personalidad. En la Tabla 13 se presenta un resumen de las medidas de personalidad incluidas en el meta-análisis.

Tabla 13. Cuestionarios de Personalidad Incluidos en el Meta-análisis

---

-	ATHURE - Test Human Relations
-	Dunnette Adjective Checklist
-	EPPS - Edwards Personal Preference Schedule
-	Escala de Adjetivos basada en Goldberg's
-	ESQ_FC - Employee Screening Questionnaire
-	FC-FFM - Medida de EF basada en los Cinco Factores
-	GPI - Gordon Personal Inventory
-	GPP - Gordon Personal Profile
-	GSDI - Ghiselli Self-Description Inventory
-	IPIP-FCM - Medida de EF basada en los Cinco Factores
-	MBTI - Myers-Briggs Type Indicator
-	NCAPS - Navy Computer Adaptive Personality Scales.
-	OPQ - Occupational Personality Questionnaire
-	PAPI - Perception and Preference Inventory
-	Pensacola Z
-	QI5F-tri - Inventario Quasi-ipsativo de Cinco Factores
-	SFCAS - Short Forced-Choice Anxiety Scale
-	SIV - Survey Interpersonal Values
-	TDOT - Thorndike Dimensions of Temperament
-	WPQ - Work Preference Questionnaire

---

Una vez aplicados todos los criterios de inclusión descritos en este apartado se obtuvo un total de 48 estudios que incluían 71 muestras independientes y la muestra total acumulada fue 10,489 sujetos. La mayor parte de estas investigaciones provienen de artículos publicados en revistas científicas. No obstante, también se han encontrado estudios en otro tipo de publicaciones. Se incluyeron en el meta-análisis datos de 41 artículos publicados, 2 tesis doctorales, 4 investigaciones no publicadas y un manual técnico.

### 3. CODIFICACIÓN DE ESTUDIOS

Una vez seleccionados los estudios que se utilizarían en el meta-análisis, se procedió a analizar y codificar la información considerada de interés para llevar a cabo los análisis cuantitativos. De cada estudio particular se obtuvo la siguiente información:

1. *Información del estudio.* En este apartado se recopila la información de los autores, año, título del documento, tipo de publicación (artículo, tesis, manual técnico, etc.) y nombre de esta.
2. *Información de la muestra.* Tamaño de la muestra, inicial y final, características de los sujetos (estudiantes, candidatos, empleados, etc.).
3. *Información del contexto experimental.* Si se trataba de un estudio experimental o correlacional.
4. *Información del diseño.* Indicando si era un diseño intra-grupo o inter-grupos.
5. *Información de la medida de personalidad.* Descripción de las características de la medida de EF utilizada: formato (ipsativo, quasi-ipsativo, normativo), número de factores y escalas que lo componen y dirección de la puntuación, si está basado o no en los cinco grandes y fiabilidad de la medida.
6. *Tamaño del efecto* expresado en  $d$  de Cohen o datos que permitiesen su cálculo, tales como el tamaño de la muestra, la media y la desviación típica en condición honesta y faking,  $r$  de Pearson,  $d$  de Glass; etc.

### 4. REGLAS DE DECISIÓN

Antes de detallar los pasos llevados a cabo para realizar los cálculos del meta-análisis, consideramos conveniente informar de que se han aplicado ciertas reglas de



decisión ante las peculiaridades de algunos de los estudios integrados en esta investigación. Estas reglas hacen referencia a tres aspectos: (1) cómo se han tratado las muestras duplicadas; (2) los resultados obtenidos de comparaciones con muestras normativas y; (3) cómo se han obtenido el tamaño del efecto de los factores de personalidad cuando las medidas no estaban basadas en el modelo de los cinco grandes.

1. *Muestras duplicadas*. La principal consecuencia de la presencia de estudios duplicados es la sobreestimación del tamaño del efecto del estudio, lo que produce, a su vez, alteraciones en otros estadísticos relevantes del meta-análisis como la varianza de error de muestreo y la desviación típica de los tamaños del efecto (Schmidt, 2008). Por este motivo, es una práctica habitual analizar la presencia de muestras no independientes y aplicar un método de ajuste sobre los datos como se ha realizado en nuestro caso. En aquellos casos en los que se detectó solapamiento entre dos o más muestras incluidas en el meta-análisis, se pudo proceder de dos modos dependiendo de las características de las investigaciones:

a) Cuando los estudios estaban duplicados, es decir compartían la misma muestra y condiciones experimentales, se incluyó la investigación más reciente excluyendo las demás réplicas.

b) En aquellos casos en los que los estudios reportaban estimadores de la relación de interés en diferentes contextos de faking con la misma muestra, se valoró las condiciones particulares de cada uno de los casos para su inclusión en el metas-análisis. De este modo, si los estudios ofrecían estimadores en un contexto de instrucciones generales de faking y bajo instrucciones específicas (por ejemplo, cometer faking que afecte a una única característica o que describa un perfil específico), sólo se incluyó en el estudio el valor de la condición general. En caso de que no se considerasen diferencias entre los diseños experimentales, se procedió a calcular el valor promedio del estimador del tamaño del efecto para su inclusión en el meta-análisis.

Estas medidas se aplicaron a los meta-análisis generales. No obstante, para los meta-análisis jerárquicos se codificaron por separado los estimadores del efecto promediados para realizar un análisis de posibles variables moderadoras, con el objetivo de obtener una medida más perfecta de la relación. Al tratarse de meta-análisis específicos, no se produce duplicidad de las muestras, aunque si



redundancia y parcial correlación en los resultados de unos y otros, es decir, no son totalmente independientes.

2. *Muestras normativas.* En los casos en los que el estudio sólo reportaba datos pertenecientes a la condición faking, para poder calcular el tamaño del efecto, como condición honesta se usaron los valores procedentes de la muestra normativa del test de personalidad utilizado en el estudio. Debido la magnitud de la muestra normativa, y tomando una perspectiva conservadora, sólo se incluyó como valor de  $n$  el número de sujetos que realizaron el estudio, es decir, se excluyó el tamaño muestral normativo. Esta estrategia también se aplicó a aquellos estudios que reportaban los estimadores del efecto fruto de la comparativa con muestras normativas.

3. *Compuestos de personalidad.* Muchos de los estudios incluidos en este meta-análisis no reportaban sus resultados basándose en el modelo de los cinco grandes factores de personalidad, sino que se componían de varias escalas que evalúan lo que podemos equiparar a facetas de la personalidad que se pueden clasificar dentro de estos cinco factores. No obstante, para poder expresar los resultados en base al modelo de los Cinco Grandes factores de personalidad (estabilidad emocional, extraversión, apertura a la experiencia, amigabilidad y conciencia) fue necesario estimar lo que comúnmente se denominan compuestos de personalidad, que son una representación de los factores de personalidad producto de la agregación de las diferentes escalas que evalúan el mismo factor. Para obtener estos compuestos se aplicaron las siguientes estrategias:

a) Cuando el estudio proporcionaba las correlaciones entre las distintas variables, es decir entre las escalas de personalidad, o se podían obtener a partir de datos normativos, se calculó la correlación del compuesto empleando las fórmulas de Schmidt y Hunter (2015). Estos autores indican que la aplicación de estas fórmulas permite una estimación más precisa del tamaño del efecto que si se calculase una correlación promedio entre las variables. La primera fórmula permite calcular la covarianza promedio ( $\bar{c}_{yy}$ ) entre los distintos indicadores:

$$\bar{c}_{yy} = \frac{1+(n-1)*\bar{r}_{yy}}{n} \quad (5)$$

Respecto a la fórmula,  $\bar{r}_{yy}$  es la correlación promedio entre las distintas variables que forman el compuesto y  $n$  es el número de variables. La segunda fórmula permite estimar la correlación para el compuesto:

$$r_{xy} = \frac{\bar{r}_{xy}}{\sqrt{\bar{c}_{yy}}} \quad (6)$$

En donde  $r_{xy}$  hace referencia al tamaño de la correlación estimada para el compuesto,  $\bar{r}_{xy}$  es la correlación promedio entre los indicadores y la variable  $x$  y  $\bar{c}_{yy}$ , tal y como hemos descrito, es la covarianza promedio entre los indicadores.

Por último, se calculó la fiabilidad del compuesto aplicando la fórmula desarrollada por Mosier (1943):

$$r_{yy} = \frac{\underline{1}'(R_{yy} - D + D_{rel})\underline{1}}{\underline{1}'R_{yy}\underline{1}} \quad (7)$$

El denominador es la varianza verdadera del compuesto y el numerador corresponde a la varianza total del compuesto.

Para aplicar estas fórmulas era necesario que los tamaños del efecto estuviesen reportados en  $r$  de Pearson, por lo tanto, se realizó una doble transformación para poder calcular los compuestos, dado que las correlaciones reportadas por los estudios primarios se estimaron en  $d$  de Cohen al tratarse de un estudio de diferencias entre condiciones de faking y no faking. Para realizar las transformaciones de los tamaños del efecto se utilizaron las fórmulas desarrolladas por Hunter y Schmidt (2004, pág. 279). Primero de  $d$  de Cohen a  $r$  de Pearson, aplicando la fórmula:

$$r = \frac{d}{\sqrt{\frac{1}{pq} + d^2}} \quad (8)$$

Una vez calculada la validez de compuesto expresada en  $r$  de Pearson, se volvió a transformar a  $d$  de Cohen para su incorporación a los meta-análisis. La fórmula utilizada fue la siguiente:

$$d = \frac{r}{\sqrt{pq(1-r^2)}} \quad (9)$$

b) En aquellos casos en los que no se pudo obtener el valor de las correlaciones entre las diferentes variables de cada compuesto, se calculó el tamaño del efecto promedio.

## 5. ANÁLISIS DE VARIABLES MODERADORAS

Los *moderadores* son aquellas variables que pueden estar afectando la relación entre otros dos indicadores (Schmidt y Hunter, 2015). En ese sentido, es de crucial importancia, en la realización de estudios meta-analíticos, tomar en consideración la posible presencia de variables moderadoras que puedan estar afectando a los resultados.

El conjunto de factores que pueden actuar de moderadores en una relación puede ser muy amplio, ya que pueden derivar de aspectos como las propias características de las variables o de las medidas aplicadas, así como de aspectos relativos a las características de los estudios como, por ejemplo, el tipo de muestra o el diseño de la investigación. Por norma general, las investigaciones primarias sobre la materia proporcionan una primera aproximación a las variables moderadoras que podrían estar afectando a los resultados y permiten generar un conjunto de moderadores a tener en cuenta para realizar los meta-análisis.

Para poder estudiar el efecto que puede tener sobre los resultados una variable moderadora, Schmidt y Hunter (2015) sugieren como método más idóneo dividir los meta-análisis en sub-grupos atendiendo a las variables que se desean analizar como moderadoras. En este estudio se han considerado dos variables moderadoras: (1) el diseño del estudio y; (2) el tipo de cuestionario de EF.

1. *Diseño del estudio*. Es la primera variable moderadora que se ha tomado en consideración a la hora de realizar los meta-análisis. En base a la literatura se han diferenciado cuatro categorías: estudios intra-grupo, inter-grupos, muestras de candidatos y muestras de empleados. Todos los estudios primarios incluidos en el meta-análisis se pudieron englobar en alguna de las categorías enumeradas.

2. *Tipo de medida de EF*. El segundo moderador que se ha analizado en este estudio es el tipo de test, distinguiendo tres categorías: test ipsativos, quasi-ipsativos y normativos. El objetivo de este estudio es determinar la resistencia al faking de los cuestionarios de EF, por lo que distinguiremos entre los tres tipos de

cuestionario para determinar si las características metodológicas de cada uno de ellos afectan de modo diferente a los resultados.

3. Por último, se ha considerado el efecto que podrían tener simultáneamente ambas variables, tipo de prueba y diseño del estudio, sobre los resultados. De tal modo que se han distinguido cuatro categorías, un ejemplo sería: ipsativo-intra-grupo, ipsativo-inter-grupos, ipsativo-candidatos e ipsativo-empleados. Esta combinación se ha repetido para cada uno de los formatos con cada uno de los factores de personalidad dando lugar a lo que Schmidt y Hunter (2015) denominan meta-análisis jerárquico, es decir, el análisis conjunto de moderadores, lo que permite un análisis más preciso de los resultados, tal y como indican estos autores. En el caso de los test normativos, no se encontraron estudios primarios suficientes para poder realizar un meta-análisis jerárquico.

## 6. ESTIMACIÓN DEL TAMAÑO DEL EFECTO

La mayor parte de los estudios primarios reportaban el tamaño del efecto expresado en  $d$  de Cohen. Estadístico en el que se expresan los resultados cuando se lleva a cabo un meta-análisis experimental de diferencias entre condiciones como en este caso, por lo que directamente se incluyeron en el meta-análisis sin requerir ningún cálculo.

Sin embargo, se encontraron estudios individuales que no reportaban un tamaño del efecto, pero proporcionaban la  $SD$ , media o, por ejemplo, un estadístico que permitía su transformación en un coeficiente de  $d$  de Cohen.

Respecto al primer caso, se procedió al cálculo del coeficiente  $d$  de Cohen con los valores de la media, la  $SD$  y el tamaño de la muestra reportados por los estudios, aplicando la fórmula de Hunter y Schmidt (2004, pág. 277). En el segundo caso, los estadísticos  $t$  y  $d$  de Glass se transformaron directamente a tamaños del efecto expresados en  $d$  de Cohen.

Por último, se comprobó la dirección de las relaciones entre las variables y, cuando fue necesario, se invirtió el signo de los tamaños del efecto para unificar las relaciones, de tal modo que el signo negativo significase que se había obtenido una puntuación menor en la condición *faking* que en la condición honesta.

## 7. FIABILIDAD

Para poder llevar a cabo las correcciones por error de medida fue necesario obtener las fiabilidades de las variables a partir de los coeficientes de los estudios primarios o, en caso de que estos no fuesen reportados en el estudio, a partir de los datos de los manuales de las medidas utilizadas. No obstante, como no fue posible obtener este valor para cada uno de los estudios primarios, se creó una distribución empírica de los coeficientes de fiabilidad para cada una de las variables analizadas a partir de los datos obtenidos. En la Tabla 14 se puede observar los valores medios de los coeficientes de fiabilidad para cada uno de los meta-análisis generales.

Todos los valores incluidos en las distribuciones son coeficientes de consistencia interna expresados en alfa de Cronbach. Ninguno de los meta-análisis realizados hasta ahora había hecho correcciones por error de medida en los resultados, por lo tanto, los valores descritos en esta tabla son una contribución única de este estudio sobre las distribuciones de las fiabilidades de las variables analizadas.

Para agregar cuantitativamente los resultados de los estudios obtenidos, se siguieron los métodos de meta-análisis experimental de diferencias desarrollados por Hunter y Schmidt (1990, 2004; Schmidt y Hunter, 2015) y se realizaron los cálculos con el software de Schmidt y Le (V2.0, 2014) que implementa esta metodología y permite corregir los efectos que los errores artificiales ejercen sobre el tamaño del efecto resultante.

Tabla 14. Distribuciones de los Coeficientes de Fiabilidad de las Variables Incluidas en los Meta-análisis

	<i>K</i>	<i>Media</i>	<i>SD</i>	<i>Mín - Máx</i>
Experimental Intra-grupo				
<i>Estabilidad emocional</i>	5	.80	.11	.63 / .87
<i>Extraversión</i>	9	.80	.09	.66 / .89
<i>Apertura a la experiencia</i>	8	.80	.09	.66 / .87
<i>Amigabilidad</i>	8	.72	.13	.49 / .83
<i>Conciencia</i>	13	.73	.10	.65 / .87
Experimental Inter-grupos				
<i>Estabilidad emocional</i>	12	.82	.08	.66 / .92
<i>Extraversión</i>	12	.82	.05	.73 / .87
<i>Apertura a la experiencia</i>	10	.80	.08	.66 / .87
<i>Amigabilidad</i>	10	.77	.06	.66 / .83
<i>Conciencia</i>	22	.72	.11	.51 / .87
Correlacional Candidatos				
<i>Estabilidad emocional</i>	3	.81	.07	.74 / .87
<i>Extraversión</i>	4	.84	.04	.78 / .85
<i>Apertura a la experiencia</i>	-	-	-	-
<i>Amigabilidad</i>	1	.78	-	.78
<i>Conciencia</i>	5	.74	.12	.55 / .84
Correlacional Empleados				
<i>Estabilidad emocional</i>	3	.82	.06	.71 / .82
<i>Extraversión</i>	4	.82	.06	.73 / .85
<i>Apertura a la experiencia</i>	2	.81	.01	.80 / .81
<i>Amigabilidad</i>	2	.75	.12	.66 / .83
<i>Conciencia</i>	7	.79	.10	.64 / .87
Ipsativos				
<i>Estabilidad emocional</i>	1	.77	-	.77
<i>Extraversión</i>	4	.74	.11	.66 / .89
<i>Apertura a la experiencia</i>	4	.71	.07	.66 / .77
<i>Amigabilidad</i>	4	.64	.12	.49 / .77
<i>Conciencia</i>	4	.74	.10	.66 / .86
Quasi-ipsativos				
<i>Estabilidad emocional</i>	16	.82	.09	.63 / .92
<i>Extraversión</i>	17	.83	.05	.73 / .87
<i>Apertura a la experiencia</i>	14	.82	.07	.66 / .87
<i>Amigabilidad</i>	14	.78	.07	.66 / .83
<i>Conciencia</i>	30	.73	.10	.51 / .87

*Nota.* *K* = número de coeficientes de fiabilidad; *Media* = valor promedio de fiabilidad de la distribución; *SD* = desviación típica de los valores de fiabilidad de la distribución; *Mín - Máx* = valores mínimo y máximo de fiabilidad de la distribución.

## RESULTADOS

A continuación, se describen los resultados obtenidos en los meta-análisis realizados. Cada una de las tablas contiene, de izquierda a derecha, el número de muestras independientes ( $K$ ), el tamaño de muestra acumulado ( $N$ ), el tamaño del efecto promedio observado ponderado por el tamaño de la muestra ( $d_w$ ), la varianza observada ( $S^2_{obs}$ ), la desviación típica observada ( $SD_{obs}$ ), la varianza debida al error de muestreo ( $S^2_e$ ), el tamaño del efecto verdadero ( $\delta$ ), la desviación típica de  $\delta$  ( $SD_\delta$ ), el porcentaje de varianza explicada por los errores artificiales (%VE), el valor mínimo de credibilidad del 90% de  $\delta$  ( $VC_\delta$  90%) y, por último, el intervalo de confianza del 95% ( $IC_\delta$  95%).

Habitualmente, los estudios meta-analíticos se llevan a cabo como un método para comprobar si se generaliza la validez. Sin embargo, en este caso se trata de estudiar si se generalizan los efectos de una conducta negativa, el faking. Por lo que la perspectiva desde la que deben ser interpretados los resultados es la opuesta, esto es, a menor generalización de las conductas de faking, mayor resistencia presentan los cuestionarios de EF a este fenómeno. Por tanto, en este punto debemos aclarar que la interpretación de los intervalos de credibilidad y de confianza varía ligeramente. Lo que se espera, dadas las hipótesis planteadas, es que los intervalos contengan el cero, lo que reflejaría la resistencia de estas medidas al faking. El intervalo de credibilidad estima la variabilidad de las correlaciones individuales a lo largo de los estudios integrados. Así, en este caso, un valor de credibilidad del 90% que incluya el cero indica que el 90% de las correlaciones no pueden considerarse estadísticamente diferentes a ese valor. Por su parte, el intervalo de confianza proporciona una estimación de la variabilidad de la correlación promedio. De este modo, un intervalo de confianza del 90% en el que el cero está incluido indica que existe un 90% de confianza de que el tamaño del efecto promedio sea igual a cero y, por tanto, indicaría que los resultados no pueden generalizarse a otros contextos (Schmidt y Hunter, 2015; Judge y Bono, 2001).

Para facilitar la comprensión de los resultados obtenidos en los meta-análisis, estos se describirán por cada factor de personalidad. En primer lugar, se presentan los resultados de los meta-análisis según el tipo de diseño (Tabla 15) y a continuación los resultados de los meta-análisis jerárquicos para cada formato de cuestionario de EF (Tablas 16 a 18).



## 1. RESULTADOS DE LOS META-ANÁLISIS SEGÚN EL DISEÑO

Los resultados obtenidos para el factor estabilidad emocional (Tabla 15) muestran, en términos generales, valores pequeños de los tamaños del efecto corregidos, lo que indica que existe una baja incidencia del faking. Sin embargo, si atendemos separadamente a los resultados obtenidos según el tipo de diseño, podemos observar que el tamaño del efecto se eleva a 0.34 ( $K = 21$ ,  $N = 2,375$ ) en los estudios experimentales intra-grupo, indicando que el tipo de diseño estaría afectando a los resultados para este factor. No obstante, si nos fijamos en los valores de los intervalos de credibilidad nos encontramos con que se incluye el cero en todos los casos. Por su parte, los intervalos de confianza incluyen o están muy próximos también a cero. Por lo tanto, estos valores muestran que estos resultados no pueden considerarse estadísticamente diferentes de cero ni generalizarse a otros contextos.

En relación con el factor extraversión, en conjunto, se observa una  $d$  corregida con valores pequeños, entre 0.15 y 0.30, la totalidad de los meta-análisis. Respecto al tipo de diseño, observamos que el tamaño del efecto es mayor, con una  $\delta = 0.30$  ( $K = 22$ ,  $N = 4,426$ ), en los estudios experimentales intra-grupo en comparación con los otros diseños, cuyos valores son inferiores a  $\delta = 0.23$ . Aunque, como en el caso del factor estabilidad emocional, estos resultados no pueden considerarse estadísticamente diferentes de cero ni generalizables a otras muestras, dado que se incluye el cero en todos los intervalos de credibilidad, con la única excepción de las muestras de empleados. Este caso el número de estudios acumulados se reduce a prácticamente la mitad, lo que podría estar condicionando este resultado. Asimismo, los valores en el caso del intervalo de confianza (valor mínimo) están muy próximos a cero.

En tercer lugar, el factor apertura a la experiencia muestra, en global, tamaños del efecto pequeños, con valores cercanos a cero en la mayoría de los meta-análisis. No obstante, destaca que, dentro de los tipos de diseños, el tamaño del efecto más alto se observe en los diseños correlacionales de empleados con una  $\delta = 0.34$  ( $K = 5$ ,  $N = 460$ ) seguido de los diseños experimentales intra-grupo con una  $d$  corregida de .26 ( $K = 22$ ,  $N = 2,122$ ). Aun y todo, este valor es muy inferior al obtenido para esta misma condición en los restantes factores. Asimismo, los valores de los intervalos de confianza y credibilidad muestran que las diferencias obtenidas no pueden considerarse relevantes ya que el cero se incluye en los intervalos o son muy próximos a este valor.

Respecto a los resultados obtenidos para el factor amigabilidad, en la mayoría de los meta-análisis los tamaños del efecto fueron negativos o cercanos a cero. Esto significa que las personas tienden a presentarse como más “rudas” en la condición faking que cuando responden honestamente. Así, atendiendo al tipo de diseño, los resultados muestran tamaños del efecto pequeños y positivos para los diseños inter-grupos ( $\delta = 0.08$ ) y de muestras de candidatos ( $\delta = 0.01$ ) y pequeños y negativos para los estudios intra-grupo ( $\delta = -0.25$ ) y de muestras de empleados ( $\delta = -0.29$ ). Si nos fijamos en el valor mínimo del intervalo de credibilidad, comprobamos que en todos los casos se incluye el cero, indicando que los tamaños del efecto obtenidos no pueden considerarse distintos de cero. Además, los valores del intervalo de confianza muestran que estos resultados no pueden generalizarse a otros contextos ya que el cero se incluye en todos los casos.

Por último, los resultados obtenidos para el factor conciencia muestran que, en términos generales, es el factor que más se ve afectado por el faking en las medidas de EF. Los tamaños del efecto oscilan entre un amplio rango de valores que van desde  $\delta = -0.25$  hasta  $\delta = 0.67$ . Centrándonos en el tipo de diseño, se observan tamaños del efecto más altos para los estudios experimentales, con una  $d$  corregida de 0.67 ( $K = 32$ ,  $N = 2,760$ ) para los diseños intra-grupo y de 0.59 ( $K = 32$ ,  $N = 5,985$ ) en el caso de los diseños inter-grupos. En cambio, en los estudios correlacionales se obtuvieron magnitudes del tamaño del efecto más pequeñas de  $\delta = 0.21$  ( $K = 13$ ,  $N = 2,432$ ) en las muestras de candidatos y de  $\delta = -0.25$  en empleados. De nuevo, se detecta que el tipo de diseño modera los efectos encontrados. No obstante, el valor del intervalo de credibilidad incluye el cero en la práctica totalidad de los casos, lo que indica que estos valores, a pesar de su magnitud, no pueden considerarse diferentes de cero. Sólo en las muestras de empleados este intervalo no contiene el cero, aunque el número de estudios acumulados para esta condición puede estar condicionando los resultados ( $K = 7$ ,  $N = 509$ ).

Tabla 15. Meta-análisis de la Resistencia al Faking de los Cuestionarios de Elección Forzosa según el Tipo de Diseño

	$K$	$N$	$d_w$	$S^2_{obs}$	$SD_{obs}$	$S^2_e$	$\delta$	$SD_\delta$	%VE	$VC_\delta$ 90%	$IC_\delta$ 95%
Experimental intra grupo											
<i>Estabilidad emocional</i>	21	2,375	0.30	.2956	.5437	.0361	0.34	.5717	12	-0.39	0.07 / 0.60
<i>Extraversión</i>	22	2,426	0.27	.1657	.4071	.0369	0.30	.4012	22	-0.22	0.11 / 0.49
<i>Apertura a la experiencia</i>	22	2,122	0.23	.2313	.4810	.0423	0.26	.4870	19	-0.36	0.03 / 0.48
<i>Amigabilidad</i>	22	2,122	-0.21	.2946	.5428	.0423	-0.25	.5917	14	0.51	-0.52 / 0.02
<i>Conciencia</i>	32	2,760	0.57	.3229	.5682	.0489	0.67	.6118	16	-0.11	0.44 / 0.90
Experimental inter grupos											
<i>Estabilidad emocional</i>	25	5,542	0.12	.1904	.4364	.0182	0.13	.4590	10	-0.46	-0.06 / 0.32
<i>Extraversión</i>	22	5,067	0.15	.0598	.2446	.0175	0.17	.2280	29	-0.12	0.06 / 0.29
<i>Apertura a la experiencia</i>	20	3,684	0.11	.1417	.3765	.0218	0.12	.3889	15	-0.37	-0.06 / 0.31
<i>Amigabilidad</i>	21	4,219	0.07	.3174	.5634	.0200	0.08	.6207	6	-0.72	-0.20 / 0.35
<i>Conciencia</i>	32	5,985	0.50	.2398	.4897	.0222	0.59	.5492	10	-0.11	0.39 / 0.79
Correlacional candidatos											
<i>Estabilidad emocional</i>	11	2,080	-0.13	.0989	.3146	.0230	-0.14	.3097	22	0.26	-0.36 / 0.07
<i>Extraversión</i>	12	2,392	0.14	.0315	.1775	.0202	0.15	.1166	64	-0.00	0.04 / 0.25
<i>Apertura a la experiencia</i>	8	1,135	-0.05	.0407	.2017	.0284	-0.05	.1109	70	0.01	-0.19 / 0.09
<i>Amigabilidad</i>	9	1,447	0.01	.1541	.3927	.0501	0.01	.4070	16	-0.51	-0.28 / 0.30
<i>Conciencia</i>	13	2,432	0.18	.0776	.2786	.0160	0.21	.2758	28	-0.14	0.04 / 0.39
Correlacional empleados											
<i>Estabilidad emocional</i>	7	986	-0.22	.0930	.3050	.0288	-0.25	.2864	31	0.11	-0.51 / 0.00
<i>Extraversión</i>	7	509	0.21	.0423	.2058	.0559	0.23	.0000	100	0.23	0.09 / 0.40
<i>Apertura a la experiencia</i>	5	460	0.30	.0930	.3049	.0443	0.34	.2460	48	0.02	0.04 / 0.64
<i>Amigabilidad</i>	6	479	-0.25	.1803	.4246	.0510	-0.29	.4169	28	0.25	-0.68 / 0.11
<i>Conciencia</i>	7	509	-0.22	.0663	.2576	.0560	-0.25	.1142	85	-0.10	-0.46 / -0.03

*Nota.* El signo negativo en las puntuaciones muestra una menor puntuación en la condición faking.  $K$  = número de muestras independientes;  $N$  = tamaño de muestra acumulado;  $d_w$  = tamaño del efecto observado ponderado por el tamaño de la muestra expresado en  $d$  de Cohen;  $S^2_{obs}$  = varianza observada;  $SD_{obs}$  = desviación típica observada;  $S^2_e$  = varianza debida al error de muestreo;  $\delta$  = tamaño del efecto verdadero expresado en  $d$  de Cohen;  $SD_\delta$  = desviación típica de  $\delta$ ; %VE = porcentaje de varianza explicada por los errores artificiales;  $VC_\delta$  90% = valor mínimo de credibilidad del 90% de  $\delta$ ;  $IC_\delta$  95% = intervalo de confianza del 95%.

## 2. RESULTADOS DE LOS META-ANÁLISIS JERÁRQUICOS SEGÚN EL TIPO DE CUESTIONARIO DE ELECCIÓN FORZOSA

En relación al tipo de test (Tablas de 16 a 18) nos encontramos con que el factor estabilidad emocional muestra valores del tamaño del efecto más elevados en los cuestionarios quasi-ipsativos en los diseños experimentales, con valores de  $\delta = 0.51$  ( $K = 12$ ,  $N = 1,610$ ) para las muestras intra-grupo y de  $\delta = 0.69$  ( $K = 8$ ,  $N = 1,937$ ) para los estudios inter-grupos. Por el contrario, en los estudios correlacionales apenas se produce faking, observándose valores de  $\delta = 0.07$  ( $K = 6$ ,  $N = 1,168$ ) cuando se trata de candidatos y de  $\delta = -0.31$  ( $K = 4$ ,  $N = 255$ ) en muestras de empleados (recordemos que el signo menos implica que se ha obtenido una menor puntuación en la condición de faking). En cuanto a los cuestionarios ipsativos los valores oscilan desde  $-0.37$  ( $K = 5$ ,  $N = 912$ ) para las muestras de empleados hasta  $0.24$  ( $K = 4$ ,  $N = 588$ ) para las muestras inter-grupos, lo que refleja una menor incidencia del faking en este formato. Respecto a los test normativos, el tamaño del efecto corregido es pequeño ( $\delta = -0.17$ ), aunque en este caso no se han podido analizar los resultados por moderadores ya que no se ha alcanzado una muestra suficiente para llevar a cabo el meta-análisis jerárquico. Respecto a los intervalos de confianza y credibilidad, se incluye el cero en todos los casos a excepción de los diseños experimentales en el caso de los tests quasi-ipsativos y del diseño inter-grupo en los cuestionarios ipsativos. No obstante, se produce error de muestreo de segundo orden en el meta-análisis de la relación entre este diseño y los tests ipsativos, por lo que se deben tratar con cautela los resultados.

Respecto al factor extraversión, los resultados no muestran importantes diferencias entre los formatos ipsativo y quasi-ipsativo, obteniéndose valores del tamaño del efecto pequeños en todos los casos. La excepción es el tamaño del efecto corregido para el formato ipsativo en la condición intra-grupo, cuyo valor es interpretado como alto según los criterios de Cohen (1989), con una  $\delta = 1.02$  ( $K = 6$ ,  $N = 366$ ). Sin embargo, si nos fijamos en los intervalos de credibilidad y confianza en todos se incluye el cero o están próximos a este valor, exceptuando, de nuevo, los correspondientes a la relación entre el formato ipsativo y la condición intra-grupo.

Apertura a la experiencia, destaca por ser el factor en el cual se han obtenidos los tamaños del efecto más altos, con valores de  $\delta = 0.43$  ( $K = 3$ ,  $N = 236$ ) en los estudios con muestras de empleados para el caso de los cuestionarios quasi-ipsativos y de  $\delta = 0.45$  ( $K = 6$ ,  $N = 366$ ) en las muestras intra-grupo para los formatos ipsativos. Estos resultados

nuevamente muestran que el diseño del estudio es un moderador de la relación entre las variables, si bien es cierto que los tamaños muestrales acumulados son pequeños, no llegando a superar los 400 sujetos en dos de los meta-análisis. En cualquier caso, los intervalos de confianza y credibilidad vuelven a incluir el cero en casi todos los análisis realizados.

En cuanto al factor amigabilidad, se observan tamaños del efecto negativos para el formato ipsativo en todos los casos, destacando el tamaño del efecto encontrado para este formato en las muestras intra-grupo  $\delta = -0.63$  ( $K = 6$ ,  $N = 366$ ) ya que es el tamaño del efecto más alto obtenido para este factor. En cambio, en los cuestionarios quasi-ipsativos, los valores más altos se corresponden con las muestras de candidatos ( $\delta = -0.50$ ) y de empleados ( $\delta = -0.41$ ). Además, se ha obtenido un tamaño del efecto corregido de 0.31 ( $K = 7$ ,  $N = 2,109$ ) para los diseños inter-grupos que aplicaron esta medida, siendo el único resultado en el que se encuentra una diferencia positiva para este factor, es decir, la única condición en la que los sujetos puntúan más alto en este factor en la condición faking. Los intervalos de confianza y credibilidad, de nuevo vuelven a incluir el cero en casi la totalidad de los meta-análisis para este factor o sus valores se aproximan a este valor.

Para el factor conciencia los resultados varían ligeramente en función del formato de EF. Es en los estudios experimentales donde se observan los tamaños del efecto más grandes. Para los quasi-ipsativos este dato se asocia a los diseños inter-grupos ( $\delta = .99$ ,  $K = 15$ ,  $N = 3,262$ ) y para los ipsativos a los diseños intra-grupo ( $\delta = 1.27$ ,  $K = 6$ ,  $N = 366$ ). Pero en este último caso la muestra y número de estudios acumulados es pequeña produciéndose error de muestreo de segundo orden. Por su parte si nos fijamos en el valor mínimo del intervalo de credibilidad, observamos que en la mayor parte de los casos los meta-análisis incluyen el cero en el intervalo, a excepción de aquellos casos en los que se ha producido error de muestreo de segundo orden y del meta-análisis para la condición inter-grupos de los cuestionarios ipsativos.

Un importante dato que se debe tener en cuenta a la hora de analizar los resultados obtenidos son los valores de la varianza y la desviación, tanto observada como corregida. Si analizados los datos hallados en los meta-análisis realizados podemos comprobar que prácticamente en la totalidad de los casos los valores de estos estadísticos, especialmente de la SD, son relevantes, siendo en gran parte de las integraciones incluso superiores al tamaño del efecto. Estos valores, por tanto, nos muestran que existe una gran variabilidad

en los resultados, incluso en los meta-análisis jerárquicos, lo que sugiere la existencia de otras variables que puedan estar afectando a los resultados.

También podemos observar que, especialmente, en el meta-análisis jerárquico de los cuestionarios de EF ipsativos se ha producido un error de muestreo de segundo orden en siete de los meta-análisis realizados debido al reducido número de estudios acumulados. Esto sugiere que se analicen los resultados obtenidos en estos meta-análisis con cautela.





Tabla 16. Meta-análisis Jerárquico de la Resistencia al Faking de los Cuestionarios de Elección Forzosa Ipsativos

	$K$	$N$	$d_w$	$S^2_{obs}$	$SD_{obs}$	$S^2_e$	$\delta$	$SD_\delta$	%VE	$VC_\delta$ 90%	$IC_\delta$ 95%
Estabilidad Emocional (EE)	13	1,958	-0.14	.1313	.3624	.0268	-0.16	.3684	20	0.31	-0.38 / 0.07
EE - Intra grupo	2	234	0.06	.0745	.2729	.0347	0.06	.1995	47	-0.20	-0.32 / 0.44
EE - Inter grupo	4	588	0.21	.0274	.1656	.0276	0.24	.0000	100	0.24	0.06 / 0.43
EE - Candidatos	5	912	-0.37	.0551	.2348	.0224	-0.37	.1809	41	-0.13	-0.57 / -0.16
EE - Empleados	2	224	-0.35	.1551	.3939	.0365	-0.35	.3445	24	0.09	-0.90 / 0.19
Extraversión (EX)	17	2,090	0.27	.1391	.3730	.0331	0.32	.3774	24	-0.17	0.11 / 0.52
EX - Intra grupo	6	366	0.87	.1231	.3508	.0729	1.02	.2512	62	0.69	0.69 / 1.34
EX - Inter grupo	4	588	0.13	.1028	.3207	.0275	0.15	.3129	27	-0.25	-0.21 / 0.51
EX - Candidatos	5	912	0.14	.0264	.1639	.0221	0.14	.0691	82	0.05	-0.00 / 0.28
EX - Empleados	2	224	0.20	.0067	.0819	.0361	0.20	.0000	100	0.20	0.08 / 0.31
Apertura a la Experiencia (AE)	17	2,090	0.05	.1926	.4389	.0328	0.06	.4730	17	-0.55	-0.19 / 0.31
AE - Intra grupo	6	366	0.38	.6656	.8156	.0678	0.45	.9266	10	-0.73	-0.33 / 1.23
AE - Inter grupo	4	588	0.01	.1167	.3417	.0274	0.01	.3406	23	-0.43	-0.38 / 0.39
AE - Candidatos	5	912	-0.09	.0272	.1649	.0221	-0.09	.0717	81	-0.00	-0.24 / 0.05
AE - Empleados	2	224	0.21	.0035	.0594	.0361	0.21	.0000	100	0.21	0.12 / 0.29
Amigabilidad (A)	17	2,090	-0.23	.2168	.4657	.0330	-0.28	.5350	15	0.40	-0.56 / -0.01
A - Intra grupo	6	366	-0.49	.7264	.8523	.0686	-0.63	1.0457	10	-0.71	-1.51 / 0.25
A - Inter grupo	4	588	-0.27	.2363	.4891	.0276	-0.31	.5206	12	0.36	-0.85 / 0.24
A - Candidatos	5	912	-0.09	.0054	.0737	.0221	-0.09	.0000	100	-0.09	-0.15 / -0.02
A - Empleados	2	224	-0.26	.0000	.0008	.0362	-0.26	.0000	100	-0.26	-0.26 / -0.26
Conciencia (C)	17	2,090	0.27	.2244	.4738	.0331	0.32	.5098	15	-0.34	0.05 / 0.58
C - Intra grupo	6	366	1.08	.0718	.2680	.0764	1.27	.0000	100	1.27	1.02 / 1.53
C - Inter grupo	4	588	0.18	.0396	.1989	.0275	0.20	.1252	70	0.04	-0.02 / 0.43
C - Candidatos	5	912	0.15	.0883	.2972	.0221	0.15	.2573	25	-0.18	-0.11 / 0.41
C - Empleados	2	224	-0.32	.0003	.0170	.0364	-0.32	.0000	100	-0.32	0.34 / -0.30

*Nota.* El signo negativo en las puntuaciones muestra una menor puntuación en la condición faking.  $K$  = número de muestras independientes;  $N$  = tamaño de muestra acumulado;  $d_w$  = tamaño del efecto observado ponderado por el tamaño de la muestra expresado en  $d$  de Cohen;  $S^2_{obs}$  = varianza observada;  $SD_{obs}$  = desviación típica observada;  $S^2_e$  = varianza debida al error de muestreo;  $\delta$  = tamaño del efecto verdadero expresado en  $d$  de Cohen;  $SD_\delta$  = desviación típica de  $\delta$ ; %VE = porcentaje de varianza explicada por los errores artificiales;  $VC_\delta$  90% = valor mínimo de credibilidad del 90% de  $\delta$ ;  $IC_\delta$  95% = intervalo de confianza del 95%.



Tabla 17. Meta-análisis Jerárquico de la Resistencia al Faking de los Cuestionarios de Elección Forzosa Quasi-Ipsativos

	$K$	$N$	$d_w$	$S^2_{obs}$	$SD_{obs}$	$S^2_e$	$\delta$	$SD_\delta$	%VE	$VC_\delta$ 90%	$IC_\delta$ 95%
Estabilidad Emocional (EE)	29	4,790	0.37	.1670	.4087	.0248	0.41	.4177	15	-0.12	0.25 / 0.58
EE - Intra grupo	12	1,610	0.45	.2314	.4810	.0309	0.51	.5037	14	-0.13	0.20 / 0.82
EE - Inter grupo	8	1,937	0.63	.0223	.1494	.0174	0.69	.0685	82	0.60	0.57 / 0.79
EE - Candidatos	6	1,168	0.06	.0528	.2299	.0207	0.07	.1994	39	-0.19	-0.14 / 0.27
EE - Empleados	4	255	-0.27	.1927	.4390	.0641	-0.31	.4053	33	0.21	-0.79 / 0.18
Extraversión (EX)	27	5,445	0.16	.0764	.2763	.0200	0.17	.2617	26	-0.16	0.06 / 0.29
EX - Intra grupo	10	1,566	0.10	.0927	.3045	.0258	0.11	.2850	28	-0.26	-0.10 / 0.32
EX - Inter grupo	6	2,426	0.18	.0834	.2889	.0100	0.20	.3013	12	-0.19	-0.06 / 0.46
EX - Candidatos	6	1,168	0.18	.0346	.1861	.0207	0.19	.1278	60	0.03	0.03 / 0.35
EX - Empleados	5	285	0.22	.0701	.2648	.0715	0.24	.0000	100	0.24	-0.02 / 0.50
Apertura a la Experiencia (AE)	25	3,716	0.20	.1526	.3906	.0273	0.23	.3912	18	-0.28	0.06 / 0.39
AE - Intra grupo	12	1,471	0.19	.1422	.3771	.0333	0.20	.3559	23	-0.25	-0.03 / 0.43
AE - Inter grupo	7	1,786	0.22	.1836	.4285	.0159	0.24	.4600	9	-0.35	-0.11 / 0.60
AE - Candidatos	3	233	0.15	.0499	.2234	.0543	0.15	.0000	100	0.15	-0.11 / 0.40
AE - Empleados	3	236	0.39	.1622	.4027	.0522	0.43	.3696	32	-0.4	-0.07 / 0.94
Amigabilidad (A)	25	4,039	0.07	.3705	.6087	.0250	0.08	.6651	7	-0.77	-0.19 / 0.35
A - Intra grupo	12	1,471	-0.05	.1315	.3626	.0331	-0.06	.3513	25	0.39	-0.29 / 0.17
A - Inter grupo	7	2,109	0.27	.4710	.6863	.0135	0.31	.7655	3	-0.67	-0.27 / 0.88
A - Candidatos	3	223	-0.50	.1048	.3237	.0559	-0.50	.2211	53	-0.22	-0.87 / -0.14
A - Empleados	3	236	-0.36	.2506	.5006	.0520	-0.41	.5166	21	0.25	-1.07 / 0.24
Conciencia (C)	46	6,615	0.60	.2495	.4995	.0294	0.70	.5479	12	-0.00	0.53 / 0.87
C - Intra grupo	20	1,900	0.62	.2818	.5309	.0448	0.73	.5706	16	-0.00	0.46 / 1.00
C - Inter grupo	15	3,262	0.83	.1378	.3713	.0201	0.99	.4039	17	0.48	0.77 / 1.22
C - Candidatos	6	1,168	0.12	.0573	.2394	.0207	0.14	.2160	36	-0.14	-0.08 / 0.35
C - Empleados	5	285	-0.14	.1045	.3233	.0713	-0.16	.2055	68	0.10	-0.48 / 0.16

*Nota.* El signo negativo en las puntuaciones muestra una menor puntuación en la condición faking.  $K$  = número de muestras independientes;  $N$  = tamaño de muestra acumulado;  $d_w$  = tamaño del efecto observado ponderado por el tamaño de la muestra expresado en  $d$  de Cohen;  $S^2_{obs}$  = varianza observada;  $SD_{obs}$  = desviación típica observada;  $S^2_e$  = varianza debida al error de muestreo;  $\delta$  = tamaño del efecto verdadero expresado en  $d$  de Cohen;  $SD_\delta$  = desviación típica de  $\delta$ ; %VE = porcentaje de varianza explicada por los errores artificiales;  $VC_\delta$  90% = valor mínimo de credibilidad del 90% de  $\delta$ ;  $IC_\delta$  95% = intervalo de confianza del 95%.

Tabla 18. Meta-análisis de la Resistencia al Faking de los Cuestionarios de Elección Forzosa Normativos

	$K$	$N$	$d_w$	$S^2_{obs}$	$SD_{obs}$	$S^2_e$	$\delta$	$SD_\delta$	%VE	$VC_\delta$ 90%	$IC_\delta$ 95%
Estabilidad Emocional	4	1,169	-0.17	.1339	.3659	.0138	-0.17	.3465	10	0.27	-0.53 / 0.19

*Nota.* El signo negativo en las puntuaciones muestra una menor puntuación en la condición faking.  $K$  = número de muestras independientes;  $N$  = tamaño de muestra acumulado;  $d_w$  = tamaño del efecto observado ponderado por el tamaño de la muestra expresado en  $d$  de Cohen;  $S^2_{obs}$  = varianza observada;  $SD_{obs}$  = desviación típica observada;  $S^2_e$  = varianza debida al error de muestreo;  $\delta$  = tamaño del efecto verdadero expresado en  $d$  de Cohen;  $SD_\delta$  = desviación típica de  $\delta$ ; %VE = porcentaje de varianza explicada por los errores artificiales;  $VC_\delta$  90% = valor mínimo de credibilidad del 90% de  $\delta$ ;  $IC_\delta$  95% = intervalo de confianza del 95%.



## DISCUSIÓN

Los resultados obtenidos en este meta-análisis demuestran que los efectos del faking no pueden considerarse estadísticamente diferentes de cero. Lo que significa que los cuestionarios de EF presentan resistencia a las conductas de faking. Por lo tanto, los resultados hallados apoyan la *H1*. Pero esta no es la única conclusión que podemos obtener de estos resultados.

Los tamaños del efecto obtenidos muestran que el faking sólo se presenta en contextos de laboratorio. Como se ha podido observar, el diseño del estudio afecta a los resultados obtenidos, encontrándose diferencias importantes en los valores del tamaño del efecto entre los diseños experimentales y los diseños correlacionales. Se muestra que el faking se produce en mayor medida en contextos de laboratorio, independientemente del tipo de medida utilizado. Estos resultados van en la misma línea de los obtenidos por Salgado (2016) para las medidas SS, al comparar muestras experimentales y correlacionales. Por lo que las *H3* y la *H3a* también obtienen apoyo empírico.

Asimismo, los valores obtenidos muestran que son los diseños intra-grupo de medidas ipsativas en los que se observan los tamaños del efecto más altos. Respecto a los diseños correlacionales, apenas se observan diferencias importantes entre las muestras de candidatos y de empleados. Aunque los tamaños muestrales pueden estar condicionando los resultados, los meta-análisis de empleados han sido realizados con una muestra acumulada menor. Estos resultados sugieren que apenas se produce faking en contextos reales de selección cuando se utilizan este tipo de medidas de personalidad. En ese sentido la *H3b* también se confirmaría.

Por otro lado, con respecto al tipo de cuestionario de EF, se puede concluir que ambos formatos permiten reducir los efectos del faking. No obstante, si nos centramos en los tamaños del efecto del factor conciencia, el más relevante para la práctica profesional, son los cuestionarios quasi-ipsativos los que han mostrado mayor resistencia a este fenómeno. Por lo que se puede confirmar parcialmente la *H2*, ya que no se obtienen los mismos resultados con cada tipo de cuestionario de EF. Sin embargo, es necesario seguir acumulando más estudios primarios para poder realizar afirmaciones más robustas, dado que en algunos de los análisis el número de estudios acumulados se reducía a 2 o 3 indicadores.

Un efecto que se produce en este tipo de medidas es el cambio en el signo del tamaño del efecto en algunos de los factores en la condición faking respecto a la condición

honesta. Esto se puede explicar debido a las características metodológicas de este tipo de pruebas (ver capítulo 5), ya que al forzar al candidato a seleccionar la opción más deseable entre varias alternativas las personas tienden a centrarse en unos factores en decremento de otros. Esto provoca que, al realizar el cómputo global para cada factor, exista una mayor diferencia entre las puntuaciones honestas y en condiciones de faking en los factores que los sujetos han considerado menos relevantes, dando lugar a un cambio en el signo en la relación.

En la realización de este estudio se han analizado dos importantes variables que la literatura científica ha identificado como moderadores de las conductas de faking: el diseño de los estudios y el tipo de cuestionario de EF utilizado. Incluso se ha llevado a cabo un meta-análisis jerárquico de estos moderadores, siendo el primero que se ha realizado sobre esta cuestión hasta el momento.

#### LIMITACIONES DEL ESTUDIO

Sin embargo, esta investigación también cuenta con ciertas limitaciones. La primera es que no se han podido llevar a cabo los análisis correspondientes al meta-análisis jerárquico de los cuestionarios normativos, al no obtenerse datos suficientes para ello. Por ese motivo, resultaría de gran interés poder ampliar este meta-análisis para poder realizar una comparativa entre todas las formas de cuestionarios de EF para llegar a conclusiones más precisas sobre su resistencia.

Asimismo, como ya hemos indicado, en algunos meta-análisis se ha producido error de muestreo de segundo orden debido al reducido número de estudios acumulados para los meta-análisis, de tal modo que los resultados podrían no reflejar el tamaño del efecto de modo preciso. En ese sentido, resulta de interés seguir aumentando el número de investigaciones para poder llegar a resultados más robustos.

Por otro lado, existen otros aspectos que no se han podido analizar en esta investigación y que también podrían actuar como variables moderadoras y afectar a los resultados. Así, por ejemplo, el tipo de empleo sobre el que basan sus respuestas en los estudios experimentales o al que optan los candidatos en los procesos reales es una variable que también es necesario analizar para saber si tiene efecto sobre los resultados.

## IMPLICACIONES PRÁCTICAS

Los resultados obtenidos han demostrado que los efectos que produce el faking cuando se utilizan medidas de personalidad de EF no es estadísticamente relevante. Esto tiene importantes implicaciones para la práctica profesional de la selección de personal. En primer lugar, los resultados muestran que los cuestionarios de EF son una herramienta útil para el control del faking, por este motivo resulta muy recomendable su uso en los procesos de selección de personal en sustitución de las medidas de personalidad SS.

En segundo lugar, en cuanto al formato del test, se puede concluir que ambos tipos de medidas de EF presentan resistencia al faking. Aunque se ha encontrado que el tipo quasi-ipsativo es el más resistente al faking, especialmente en el factor conciencia. Por lo tanto, se recomienda el uso de este tipo de cuestionarios en la práctica profesional.

## SUGERENCIAS DE INVESTIGACIÓN

En lo que respecta a las futuras investigaciones, se recomienda la realización de más estudios primarios para obtener resultados más robustos sobre variables que no han podido incluirse en este meta-análisis o de las que apenas se han podido encontrar estudios como, por ejemplo, investigaciones que analicen la resistencia al faking de los cuestionarios normativos en contextos tanto experimentales como de campo.

Asimismo, también es necesario llevar a cabo estudios primarios que pongan su atención en otras variables que no se han incluido en este estudio pero que pueden afectar a la incidencia del faking. Por ejemplo: (a) el tipo de empleo, ya que la literatura ha mostrado que los sujetos pueden modificar sus respuestas en función del puesto ofertado (Sackett et al., 1989) o (b) las características personales de los sujetos, ya que las diversas teorías de los antecedentes del faking plantean que las capacidades de cada sujeto también afectan a sus conductas de faking (McFarland y Ryan, 2000; Paulhus, 1991). Así como otras variables que sería necesario seguir investigando para obtener resultados más robustos sobre las conductas de faking.



## ESTUDIO 2

EFFECTOS DEL FAKING SOBRE LA ESTRUCTURA  
FACTORIAL DE LOS CUESTIONARIOS DE  
PERSONALIDAD DE ELECCIÓN FORZOSA QUASI-  
IPSATIVOS: INVARIANZA/EQUIVALENCIA DE LA  
MEDIDA





## INTRODUCCIÓN

El principal objetivo de este estudio es analizar si el faking afecta a la estructura factorial de los cuestionarios de personalidad de EF, en otras palabras, establecer la invarianza o equivalencia de la medida (Millsap, 2011). Salgado (2016) planteó, en un reciente meta-análisis sobre los efectos del faking en las medidas SS, que el faking afecta a la validez de constructo de estas medidas de personalidad. Señaló que, cuando el faking es univariante, éste provoca una reducción de las cargas factoriales de los componentes del constructo, mientras que cuando el faking es multivariante, éste produce factores adicionales. Por tanto, los efectos del faking sobre los cuestionarios SS se traducen en el desacuerdo en la medida (falta de equivalencia) en condiciones diferentes (honesta/faking).

Estudios primarios que han analizado el efecto del faking sobre la validez de constructo de medidas de personalidad SS, por ejemplo, Schmidt y Ryan (1993) y Celler et al. (1996), mostraron el faking producía factores adicionales cuando los participantes estaban motivados a distorsionar sus respuestas. Por su parte, Ellingson et al. (1999) y Van Iddekinge et al. (2005) encontraron que el faking producía una reducción de factores. Por tanto, estos estudios muestran que la estructura factorial de las medidas de personalidad SS podría verse afectada por el faking.

Los estudios que se han publicado sobre la validez de constructo de los cuestionarios de personalidad se han centrado en determinar si la estructura factorial de las medidas de personalidad SS se mantiene cuando se utiliza un formato de EF, encontrándose en todos los casos que la misma no se ve afectada (Cattell y Brennan, 1994; Lee et al., 2018). Sin embargo, hasta el momento, no se ha analizado qué efectos tiene el faking sobre la invarianza o equivalencia de la medida de los cuestionarios de personalidad de EF, esto es, sobre la validez de constructo.

El meta-análisis presentado en el estudio 1 de esta tesis ha mostrado la resistencia de los cuestionarios de EF a los efectos del faking, habiéndose encontrado tamaños del efecto muy semejantes entre la condición honesta y la condición faking.

Por todo ello, es probable que el faking no afecte a la estructura factorial, en otras palabras, que exista invarianza o equivalencia de la medida. En ese sentido, se propone la siguiente hipótesis:

*H1: La estructura factorial de las medidas de personalidad de EF es semejante/equivalente en las condiciones honesta y faking.*

Por otro lado, en el estudio 1 también se ha mostrado que el diseño del estudio afecta a la ocurrencia de faking. Los resultados del meta-análisis indican que el faking presenta un tamaño del efecto pequeño en todas las condiciones, aunque la magnitud se incrementa en contextos de laboratorio con diseños intra-grupo, a diferencia de lo que ocurre con los cuestionarios SS, en los que las diferencias son muy apreciables (Ones y Viswesvaran, 1999). Por ello, consideramos relevante analizar cómo afecta el faking a la estructura factorial cuando se emplean diseños entre e intra-grupos. Se plantea la siguiente hipótesis:

*H2: La estructura factorial de las medidas de personalidad de EF no se ve afectada por el tipo de diseño experimental empleado (entre o intra-grupo).*

## MÉTODO

### 1. MUESTRA

La muestra se compuso de 1,141 sujetos de los cuales el 67% eran mujeres ( $N = 764$ ) y el 32.86 % ( $N = 375$ ) hombres. La edad promedio fue de 21.05 años ( $SD = 4.04$ ) y el 49.65% tenía entre 17 y 20 años.

Todos los participantes eran estudiantes de la Universidad de Santiago de Compostela. El 97.46% estaba cursando o había finalizado recientemente una titulación de primer ciclo, mientras que sólo un 1.93% procedía de una titulación de segundo ciclo y menos del 1% cursaba estudios de doctorado. Respecto a las titulaciones de procedencia, más de la mitad de la muestra (56.13%) procedía de titulaciones del área de las ciencias sociales, seguida del área de ciencias de la salud (27.15%). En la Tabla 19 se describen en mayor detalle las características de la muestra.

Para el reclutamiento de la muestra se siguieron dos estrategias: (1) el reclutamiento de estudiantes en las aulas ( $N = 384$ , 33.65%) y; (2) la participación voluntaria en el estudio experimental ( $N = 757$ , 66.35%). En el segundo caso, para incentivar la participación de los sujetos, se anunció la realización del estudio mediante anuncios en las diferentes facultades de la Universidad de Santiago de Compostela, en los que se informaba que cada persona recibiría una compensación económica por participar voluntariamente en una investigación científica. Para poder participar en el estudio se exigió que los participantes cumpliesen la condición de ser estudiante universitario.

Tabla 19. Características Demográficas de la Muestra ( $N = 1,141$ )

	<i>N</i>	%
Sexo		
<i>Mujeres</i>	766	67.13
<i>Hombres</i>	375	32.86
Edad		
<i>17-20</i>	567	49.69
<i>20-25</i>	471	41.28
<i>≥25</i>	98	8.59
<i>No disponible</i>	5	0.44
Nivel de estudios		
<i>Primer ciclo</i>	1112	97.46
<i>Segundo ciclo</i>	22	1.93
<i>Tercer ciclo</i>	7	0.61
Área de estudio		
<i>Ciencias sociales</i>	647	56.71
<i>Ciencias</i>	101	8.85
<i>Ciencias de la salud</i>	310	27.17
<i>Humanidades</i>	54	4.73
<i>Ingenierías</i>	29	2.54

*Nota.* *N* = número de sujetos; Primer ciclo = estudios de diplomatura, licenciatura o grado; Segundo ciclo = estudios de máster; Tercer ciclo = estudios de doctorado.

Todos los participantes firmaron un documento de consentimiento informado y se garantizó la confidencialidad de la información, su uso para fines exclusivamente de la investigación llevada a cabo y el anonimato en el tratamiento de los datos.

Respecto al diseño experimental, el 43% ( $N = 490$ ) de la muestra participó en un diseño intra-grupo en el que los mismos participantes contestaron al cuestionario de modo honesto y bajo condiciones que inducían a cometer faking. El restante 57% participó en un diseño de dos grupos independientes en el que 449 sujetos cubrieron el cuestionario en la condición honesta y 202 participantes bajo instrucciones de faking.

## 2. INSTRUMENTO DE MEDIDA DE LA PERSONALIDAD

*QI5F\_tri*. La personalidad fue evaluada con el cuestionario quasi-ipsativo de EF QI5F\_tri, desarrollado por Salgado (2014), teniendo como referencia teórica el Modelo de los Cinco Grandes Factores de personalidad. Este cuestionario está compuesto por 140 tríadas que evalúan los Cinco Grandes factores de personalidad. Cada uno de los Big Five se evalúa mediante 28 ítems y cada ítem contiene tres alternativas de respuesta, una de ellas perteneciente al factor que se pretende evaluar y otras dos pertenecientes a un factor

diferente. Por tanto, para cada factor hay 28 tríadas, lo que hace que el número total de alternativas sea  $5 \times 28 \times 3 = 420$ . Todas las alternativas se presentan en forma positiva, es decir, no hay alternativas negativas. El sujeto debe responder cuál de las alternativas prefiere o le gusta más y cuál es la que le gusta menos o lo describe peor. Por tanto, la persona evaluada debe señalar dos respuestas en cada ítem. Las alternativas de respuesta pueden ser breves sentencias o adjetivos, con una puntuación semejante en deseabilidad social. La selección de las alternativas se basó en un estudio con ocho expertos en selección de personal (todos ellos doctores en psicología) que evaluaron más de 600 alternativas de acuerdo a su grado de deseabilidad social. Esta última característica es importante porque el principal objetivo del cuestionario es ser utilizado en procesos de selección de personal. Por ejemplo, un ítem puede incluir una frase que evalúa la apertura a la experiencia, otra que evalúa estabilidad emocional y otra que evalúa conciencia. Ningún ítem contiene dos frases o adjetivos que evalúan el mismo factor de personalidad y ningún ítem se utiliza para medir dos o más factores de personalidad. De este modo, la estrategia de Horn (1971) fue implementada y la puntuación para cada uno de los cinco grandes factores es algebraicamente independiente de la puntuación en los otros factores de personalidad. Un ejemplo de un ítem de este cuestionario es el siguiente:

*Soy una persona:*

- a) poco nerviosa en general*
- b) que trabajo bien con todo el mundo*
- c) comprensiva con los demás*

Por lo que se refiere al método de puntuación, el QI5F\_tri emplea el siguiente sistema. Para cada ítem, si quien contesta la alternativa correspondiente al factor marca la opción “la que mejor me describe”, se le asignan 2 puntos a ese ítem. Si la alternativa correspondiente al factor se marca como la que “menos me describe”, se le asigna 0 puntos al ítem. Si la persona no marca ninguna de las dos opciones para la alternativa del factor, entonces se le asigna 1 punto para este ítem. Por tanto, para cada uno de los Cinco Grandes, el rango potencial de puntuaciones puede oscilar entre 0 y 56 puntos. Este sistema de puntuación sigue el modelo empleado en la última versión del Gordon Personal Profile Inventory (GPP-I) de Gordon (1993).

El manual técnico del QI5F\_tri indica que los coeficientes de consistencia interna para estabilidad emocional (EE), extraversión (EX), apertura a la experiencia (AE), amigabilidad (A) y conciencia (C) fueron .71, .73, .80, .66 y .80, respectivamente. El

manual también indica que las fiabilidades test-retest para un intervalo de cuatro semanas fueron .91, .90, .79, .65 y .72 para EE, EX, AE, A y C, respectivamente. El manual técnico también incluye evidencia de la validez convergente obtenida empleando un inventario de personalidad SS. Las correlaciones observadas entre las escalas que medían el mismo factor fueron .65, .73, .60, .70 y .70 para EE, EX, AE, A y C, respectivamente. Estas correlaciones corregidas por error de medida fueron .84, .91, .73, .98 y .85 para EE, EX, AE, A y C.

### 3. PROCEDIMIENTO

Este cuestionario de EF se utilizó bajo dos condiciones experimentales: honesta y faking. En la condición honesta, los participantes siguieron las instrucciones propias del cuestionario que a continuación se describen:

*“En el siguiente cuestionario se le presentarán conjuntos de frases agrupadas en tríos. Trate de clasificarlas identificando primero la que mejor le describa, la que lo hace en segundo lugar y, finalmente, la que menos le describa. Marque un signo más (+) al lado de la frase que mejor le describa y un signo menos (–) al lado de la frase que menos le describa. Deberá dejar en blanco la que ha considerado en segundo lugar”.*

Para la condición faking, se modificaron ligeramente las instrucciones del test de tal modo que los participantes pudiesen incurrir en faking. Las instrucciones fueron las siguientes:

*“A continuación, se le presentarán conjuntos de frases agrupadas en tríos. Trate de clasificarlas identificando primero la que mejor le describa, la que lo hace en segundo lugar y, finalmente, la que menos le describa. Al responder imagínese que está en la última parte del proceso de selección para un puesto muy atractivo, dado que le ofrece grandes posibilidades de desarrollar su carrera profesional. Usted quiere conseguir ese puesto. Para ello, debe contestar a este test tratando de dar la mejor imagen posible para conseguir ese puesto. Marque un signo más (+) al lado de la frase que mejor le describa y un signo menos (–) al lado de la frase que menos le describa. Deberá dejar en blanco la que ha considerado en segundo lugar”.*

## RESULTADOS

Para examinar si el formato de EF del cuestionario QI5F\_tri es resistente al faking, es decir, si existe invarianza de medida en la estructura del cuestionario, se realizaron una serie de análisis factoriales exploratorios (AFE) y análisis factoriales confirmatorios (CFA) con las respuestas al cuestionario en condiciones de respuesta honesta y de respuesta en condiciones de faking.

En el caso del análisis factorial exploratorio se realizó empleando el método de máxima verosimilitud (*Maximum Likelihood*) porque permite la obtención de índices de ajuste absolutos, relativos y testar la equivalencia del modelo. También se realizó el análisis empleando el método de componentes principales. En todos los casos se utilizó la matriz de correlaciones de Pearson, se retuvieron cinco factores y se empleó la rotación Varimax para obtener factores ortogonales. Para llevar a cabo estos análisis se ha empleado el programa FACTOR de análisis factorial exploratorio (no restringido) de Lorenzo-Seva y Ferrando (2018), empleando la última versión del mismo.

A modo de examen complementario, los análisis factoriales se replicaron empleando como método el análisis de componentes principales, utilizando también el programa FACTOR que proporciona estadísticos de ajuste del modelo (p. ej., RMSR).

Por lo que respecta a los análisis factoriales confirmatorios, se utilizó el programa Lisrel 8 (Jöreskog y Sörbom, 1998), empleando el método de máxima verosimilitud, para comprobar si los datos se ajustaban a la estructura factorial propuesta.

Por último, se examinaron los resultados de los coeficientes de congruencia factorial (coeficientes de Burt-Tucker) para establecer el grado de congruencia entre los factores que miden el mismo constructo en condiciones de respuesta honesta y faking y el grado de incongruencia con respecto a los factores divergentes (que miden diferentes constructos).



Tabla 20. Agrupación de los Ítems del Cuestionario de EF en Clústers

Factor	Variable	Ítems
EE	V1	1ee-7ee
	V2	8ee-14ee
	V3	15ee-21ee
	V4	22ee-28ee
EX	V5	1ex-7ex
	V6	8ex-14ex
	V7	15ex-21ex
	V8	22ex-28ex
AE	V9	1ae-7ae
	V10	8ae-14ae
	V11	15ae-21ae
	V12	22ae-28ae
A	V13	1a-7a
	V14	8a-14a
	V15	15a-21a
	V16	22a-28a
C	V17	1c-7c
	V18	8c-14c
	V19	15c-21c
	V20	22c-28c

*Nota.* EE = estabilidad emocional; EX = extraversión; AE = apertura a la experiencia; A = amigabilidad; C = conciencia.

Para el propósito de este estudio, los 140 ítems que componen el cuestionario QI5F\_tri se agruparon de modo arbitrario (no se hizo aleatoriamente en el sentido estadístico estricto), en 20 clusters o compuestos (a lo largo de este estudio utilizaremos ambos términos indistintamente) cada uno de los cuales incluía 7 preguntas del cuestionario que evalúan el mismo factor de personalidad. Así, los 28 ítems que evalúan cada factor de personalidad se dividió en cuatro compuestos de siete ítems. Este procedimiento de agrupamiento de ítems se empleó tanto para los datos obtenidos en las dos condiciones experimentales (honesta y faking) de las muestras analizadas (total, inter e intra-grupo). En la Tabla 20 se indican los clusters/compuestos creados.

## 1. ANÁLISIS FACTORIAL EXPLORATORIO DE LA ESTRUCTURA DEL QI5F\_TRI

Como se ha señalado en las páginas precedentes, se ha realizado una serie de análisis factoriales exploratorios con el método de máxima verosimilitud con los datos

correspondientes a las condiciones honesta y faking, obtenidos tanto con un diseño entre grupos como con un diseño intra-grupo.

### 1.1 Análisis factorial exploratorio en la muestra total

El primer EFA se efectuó con la muestra total de sujetos que contestaron el Q15f\_tri en con instrucciones de responder honestamente, ya que los resultados de dicho análisis servirán, además de para establecer la estructura del cuestionario, para poder compararlos con los realizados con las respuestas dadas bajo instrucciones de faking. En la Tabla 21 se presentan los estadísticos descriptivos para cada una de las variables de la muestra total.

Tabla 21. Estadísticos Descriptivos de las Variables para la Muestra Total

Variable	Honesto ( $N = 939$ )		Faking ( $N = 692$ )	
	Media	SD	Media	SD
V1	5.74	5.93	5.70	4.30
V2	5.77	5.89	6.83	5.05
V3	5.07	5.30	5.20	3.86
V4	6.07	4.97	7.09	4.59
V5	7.03	7.76	5.84	5.64
V6	7.31	6.96	6.75	4.74
V7	6.39	5.30	6.65	4.22
V8	7.25	6.32	6.87	4.96
V9	7.06	8.16	8.08	6.53
V10	7.40	7.06	7.56	4.69
V11	6.83	9.37	7.88	6.94
V12	7.20	9.17	7.08	7.64
V13	8.33	7.08	5.99	6.42
V14	7.38	5.61	6.34	5.12
V15	6.36	4.10	7.35	3.86
V16	7.50	5.47	7.10	4.63
V17	6.10	7.35	6.89	6.44
V18	6.52	6.01	7.84	6.17
V19	7.34	9.20	9.26	7.57
V20	6.41	6.14	8.44	6.94

En el conjunto de los 20 clusters, se encontraron puntuaciones medias más altas y desviaciones típicas más bajas en la condición faking que en la condición honesta, aunque la diferencia para la muestra global es pequeña ya que el promedio de las medias es de

6.75 ( $SD = 6.66$ ) en la condición honesta y de 7.04 ( $SD = 5.52$ ) en la condición faking. La  $d$  de Cohen para estos resultados globales es de tan solo .04, lo que indica que el efecto global del faking es de muy escasa magnitud, aunque los efectos sobre los compuestos pueden ser de mayor relevancia.

La Tabla 22 muestra la matriz de cargas rotadas (las matrices de correlaciones factorizadas de la muestra total se pueden consultar en los Anexos 2 y 3). Todas las variables tienen su carga más alta y positiva en el factor en el que tendrían que cargar. Así, los cuatro compuestos de estabilidad emocional cargan juntos en el factor 3, los cuatro correspondientes a extraversión cargan juntos en el factor 1, los cuatro compuestos de apertura a la experiencia cargan juntos en el factor 4, los cuatro de amigabilidad cargan juntos en el factor 2 y los cuatro compuestos de conciencia cargan juntos en el factor 5. Además, puede observarse que ninguno de los factores tiene cargas positivas relevantes y que las cargas negativas son todas ellas de pequeña magnitud.

Tabla 22. Matriz de Cargas Factorial Rotadas del Análisis Factorial Exploratorio del Q15F\_tri Contestado Bajo Instrucciones de Honestidad, Empleando la Muestra Total ( $N = 939$ )

Variable	F1	F2	F3	F4	F5
<i>V1</i>	-0.080	-0.191	<b>0.520</b>	-0.139	-0.084
<i>V2</i>	0.180	-0.107	<b>0.559</b>	-0.045	-0.213
<i>V3</i>	-0.086	0.021	<b>0.519</b>	-0.066	-0.265
<i>V4</i>	-0.203	0.010	<b>0.610</b>	-0.086	0.012
<i>V5</i>	<b>0.727</b>	-0.148	-0.095	-0.104	-0.137
<i>V6</i>	<b>0.680</b>	-0.148	-0.026	-0.063	-0.272
<i>V7</i>	<b>0.506</b>	0.026	-0.097	-0.032	-0.191
<i>V8</i>	<b>0.712</b>	-0.093	0.007	-0.142	-0.134
<i>V9</i>	-0.116	-0.140	-0.063	<b>0.702</b>	-0.197
<i>V10</i>	-0.091	-0.126	0.003	<b>0.753</b>	-0.053
<i>V11</i>	0.002	-0.096	-0.035	<b>0.771</b>	0.040
<i>V12</i>	-0.050	0.091	-0.063	<b>0.774</b>	-0.103
<i>V13</i>	-0.196	<b>0.653</b>	-0.021	-0.123	-0.213
<i>V14</i>	-0.116	<b>0.658</b>	-0.192	-0.121	-0.039
<i>V15</i>	-0.107	<b>0.378</b>	0.238	-0.023	-0.051
<i>V16</i>	-0.065	<b>0.631</b>	0.044	-0.084	-0.061
<i>V17</i>	<b>-0.407</b>	-0.103	-0.123	-0.129	<b>0.401</b>
<i>V18</i>	-0.332	-0.136	-0.140	-0.168	<b>0.537</b>
<i>V19</i>	-0.207	-0.097	-0.101	-0.130	<b>0.798</b>
<i>V20</i>	-0.190	-0.059	-0.009	-0.047	<b>0.615</b>

En la Tabla 23 se presentan los estadísticos de bondad de ajuste del modelo de cinco factores ortogonales. Puede observarse que todos los estimadores muestran un excelente ajuste del modelo. Tanto los estadísticos de ajuste absoluto (p. ej., RMSR), como de ajuste relativo (p. ej., RMSEA, CFI, NNFI) como de equivalencia tienen una magnitud que se sitúa entre los puntos de corte típicos de ajuste buenos y excelentes. Por ejemplo, la RMSEA es .031, el índice NNFI es .979, el CFI es .989, y el tamaño de estadístico T para comprobar la significatividad de los valores de RMSEA y CFI indica que son estadísticamente significativos. Además, el indicador de ajuste absoluto RMSR es .0177, frente a un valor esperado de .0327, es decir un 54.1% más pequeño.

Tabla 23. Indicadores de Bondad de Ajuste del Modelo de Cinco Factores con Instrucciones de Respuesta Honesta al QI5F\_Tri, Empleando la Muestra Total ( $N = 939$ )

---

Root Mean Square Error of Approximation (RMSEA) = 0.031
Estimated Non-Centrality Parameter (NCP) = 234.500
Degrees of Freedom = 100
Test of Approximate Fit $H_0$ : RMSEA < 0.05 ( $P = 1.000$ )
Normal Theory Chi Square with 100 degrees of freedom = 189.659 ( $P = 0.000010$ )
Chi-Square for independence model with 190 degrees of freedom = 8432.496
Non-Normed Fit Index (NNFI; Tucker & Lewis) = 0.979
Comparative Fit Index (CFI) = 0.989
Schwarz's Bayesian Information Criterion (BIC) = 1011.037
Goodness of Fit Index (GFI) = 0.997
Adjusted Goodness of Fit Index (AGFI) = 0.994
Goodness of Fit Index without diagonal values (GFI) = 0.994
Adjusted Goodness of Fit Index without diagonal values (AGFI) = 0.988
T-Size Root Mean Square Error of Approximation (Ts-RMSEA) = 0.038
Maximum misspecification (T-size) = 0.141
T-Size Comparative Fit Index (Ts-CFI) = 0.982
Maximum misspecification (T-size) = 0.151
Root Mean Square of Residuals (RMSR) = 0.0177
Kelley's Expected mean value of RMSR for an acceptable model = 0.0327

---

El segundo EFA (Tabla 24) se hizo con las respuestas al QI5F\_tri bajo instrucciones que inducían faking. Como en el caso anterior, la estructura de cinco factores ortogonales fue apoyada por los datos, aunque existen ligeras diferencias. Puede apreciarse que 18 de las 20 variables factorizadas tienen su carga más alta y positiva en el factor en el que tendrían que cargar. Por ejemplo, los cuatro compuestos de estabilidad emocional cargan juntos en el factor 4, los cuatro correspondientes a extraversión cargan juntos en el factor 3, los cuatro compuestos de apertura cargan juntos en el factor 1, y los

cuatro de conciencia cargan juntos en el factor 2. Sin embargo, solo dos de los compuestos de amigabilidad (V15 y V16) tienen una carga relevante en el factor correspondiente y los otros dos tienen una carga negativa en el factor de conciencia. Además, puede observarse que ninguno de los factores tiene cargas positivas adicionales relevantes y que las cargas negativas son todas ellas de pequeña magnitud, a excepción de las dos mencionadas de amigabilidad.

Tabla 24. Matriz de Cargas Factoriales Rotadas del Análisis Factorial Exploratorio del Q15F\_tri Contestado Bajo Instrucciones de Faking, Empleando la Muestra Total ( $N = 692$ )

Variable	F1	F2	F3	F4	F5
V1	-0.174	-0.065	-0.061	<b>0.426</b>	-0.025
V2	0.034	0.036	0.107	<b>0.509</b>	0.160
V3	-0.104	-0.237	-0.031	<b>0.445</b>	-0.004
V4	-0.004	0.242	-0.182	<b>0.454</b>	0.045
V5	-0.070	-0.189	<b>0.539</b>	-0.068	-0.173
V6	-0.099	-0.245	<b>0.584</b>	0.030	-0.073
V7	-0.039	-0.018	<b>0.477</b>	0.031	0.204
V8	-0.118	-0.152	<b>0.579</b>	-0.006	0.011
V9	<b>0.605</b>	0.095	-0.007	0.005	0.048
V10	<b>0.642</b>	0.054	-0.039	0.060	-0.076
V11	<b>0.681</b>	0.161	-0.005	0.020	-0.016
V12	<b>0.630</b>	-0.247	-0.126	-0.148	0.031
V13	-0.274	<b>-0.601</b>	-0.202	-0.285	<b>0.013</b>
V14	-0.197	<b>-0.404</b>	-0.229	-0.277	<b>0.192</b>
V15	-0.002	0.072	-0.047	0.088	<b>0.605</b>
V16	-0.137	-0.317	-0.158	-0.180	<b>0.272</b>
V17	-0.091	<b>0.479</b>	-0.216	-0.030	-0.063
V18	-0.019	<b>0.661</b>	-0.205	-0.009	0.010
V19	-0.013	<b>0.702</b>	-0.194	-0.089	0.066
V20	-0.001	<b>0.627</b>	-0.190	-0.017	0.028

Por lo que se refiere a los índices de ajuste del modelo de cinco factores en el caso de las respuestas con instrucciones de faking, los mismos se presentan en la Tabla 25. Puede observarse que todos los estimadores muestran un excelente ajuste del modelo y que son prácticamente idénticos a los obtenidos en condiciones de respuesta honesta, ya que las variaciones en los estadísticos se dan en las milésimas. Así, los estadísticos de ajuste absoluto (p. ej., RMSR), de ajuste relativo (p. ej., RMSEA, CFI, NNFI) y de equivalencia tiene una magnitud que se sitúa entre los puntos de corte típicos de ajuste buenos y excelentes. Por ejemplo, la RMSEA es .026 (ligeramente mejor que en

condiciones de respuesta honesta), el índice NNFI es .977 (vs. .979 en respuesta honesta), el CFI es .988 (vs. .989 en respuesta honesta), y el tamaño de estadístico T para comprobar la significatividad de los valores de RMSEA y CFI indica que son estadísticamente significativos (.035 vs. .038 y .976 vs. .982 para instrucciones de faking y honesta, respectivamente). Además, el indicador de ajuste absoluto RMSR es .0209, frente a un valor esperado de .0380, es decir un 55% más pequeño.

Tabla 25. Indicadores de Bondad de Ajuste del Modelo de Cinco Factores con Instrucciones de Respuesta que Inducían Faking al QI5F/tri Empleando la Muestra Total ( $N = 692$ )

---

Root Mean Square Error of Approximation (RMSEA) =	0.026
Estimated Non-Centrality Parameter (NCP) =	172.750
Degrees of Freedom =	100
Test of Approximate Fit $H_0$ : RMSEA < 0.05 (P =	1.000)
Normal Theory Chi Square with 100 degrees of freedom =	147.356 (P = 0.0015)
Chi-Square for independence model with 190 degrees of freedom =	4168.723
Non-Normed Fit Index (NNFI; Tucker & Lewis) =	0.977
Comparative Fit Index (CFI) =	0.988
Schwarz's Bayesian Information Criterion (BIC) =	932.106
Goodness of Fit Index (GFI) =	0.995
Adjusted Goodness of Fit Index (AGFI) =	0.990
Goodness of Fit Index without diagonal values (GFI) =	0.986
Adjusted Goodness of Fit Index without diagonal values (AGFI) =	0.974
T-Size Root Mean Square Error of Approximation (Ts-RMSEA) =	0.035
Maximum misspecification (T-size) =	0.121
T-Size Comparative Fit Index (Ts-CFI) =	0.976
Maximum misspecification (T-size) =	0.132
Root Mean Square of Residuals (RMSR) =	0.0209
Kelley's Expected mean value of RMSR for an acceptable model =	0.0380

---

### *1.2 Análisis factorial exploratorio en el diseño intra-grupo*

Los dos siguientes EFA que se presentan fueron realizados con las respuestas bajo instrucciones de responder con honestidad e induciendo faking, pero empleando un diseño intra-grupo. Por tanto, estos dos EFA se realizaron con una submuestra de la muestra total.

Como ya hemos indicado repetidamente en esta tesis, los resultados empíricos del estudio meta-analítico de Ones y Viswesvaran (1999) para cuestionarios single-stimulus y el estudio 1 de esta tesis, muestran que el tamaño del efecto del faking se produce en

mayor medida en diseños experimentales intra-grupo. Por ello, a continuación, analizaremos específicamente como afecta el faking a la varianza de constructo en una muestra intra-grupo cuando se usan medidas de EF para examinar si se producen diferencias significativas respecto a los resultados obtenidos para la muestra total.

En la Tabla 26 se presentan los estadísticos descriptivos de cada una de las variables para la muestra intra-grupo.

Tabla 26. Estadísticos Descriptivos de las Variables en el Diseño Intra-Grupo ( $N = 490$ )

Variable	Honesta		Faking	
	Media	SD	Media	SD
V1	5.57	6.07	5.56	4.33
V2	5.74	5.93	6.69	4.96
V3	5.09	5.09	5.08	3.79
V4	6.10	4.60	7.14	4.83
V5	6.74	7.62	5.78	5.99
V6	7.32	6.97	6.62	4.99
V7	6.29	5.08	6.58	4.11
V8	6.98	6.26	6.80	5.13
V9	7.43	8.05	8.13	6.10
V10	7.60	7.11	7.59	4.74
V11	7.16	8.47	7.97	7.15
V12	7.79	8.81	7.25	7.57
V13	8.32	7.29	6.08	6.95
V14	7.31	5.19	6.38	5.16
V15	6.36	4.10	7.31	3.73
V16	7.48	5.12	7.25	4.62
V17	6.14	6.97	7.02	6.92
V18	6.49	6.57	8.05	6.61
V19	7.22	9.61	9.35	8.18
V20	6.49	6.41	8.58	7.23

Los estadísticos descriptivos muestran que, en conjunto, las medias en la condición faking son más altas que en la condición honesta, con valores promedio de 7.06 y 6.78, respectivamente. Lo opuesto ocurre con la desviación típica, los valores tienden a ser más pequeños para la condición faking, obteniéndose una SD promedio de 5.56 mientras que para la condición honesta es de 6.57. Estos resultados concuerdan con los presentados previamente. A continuación, analizaremos si también afecta a la estructura factorial de la personalidad cuando se usan cuestionarios de EF en muestras intra-grupo.



Se pueden consultar las matrices de correlaciones factorizadas del diseño intra-grupo en el apartado Anexos 4 y 5.

En la Tabla 27 se muestran la matriz de cargas factoriales en la condición honesta. Puede observarse que los resultados replican la estructura de cargas mostrada para la muestra total con notable similitud y, por tanto, los comentarios efectuados en el caso de la muestra total se aplican en su totalidad en el presente caso. Así pues, la estructura de cinco factores ha sido apoyada por los datos.

Tabla 27. Matriz de Cargas Factoriales Rotadas del Análisis Factorial Exploratorio del Q15F\_tri Contestado Bajo Instrucciones de Honestidad, con un Diseño Intra-Grupo ( $N = 490$ )

Variable	F1	F2	F3	F4	F5
<i>V1</i>	-0.137	<b>0.485</b>	-0.112	-0.123	-0.175
<i>V2</i>	-0.070	<b>0.572</b>	-0.229	0.173	-0.115
<i>V3</i>	-0.064	<b>0.516</b>	-0.253	-0.084	-0.004
<i>V4</i>	-0.115	<b>0.589</b>	0.028	-0.206	0.025
<i>V5</i>	-0.115	-0.109	-0.127	<b>0.740</b>	-0.135
<i>V6</i>	-0.082	-0.012	-0.285	<b>0.638</b>	-0.152
<i>V7</i>	-0.008	-0.085	-0.173	<b>0.529</b>	0.003
<i>V8</i>	-0.117	0.010	-0.143	<b>0.702</b>	-0.076
<i>V9</i>	<b>0.692</b>	-0.088	-0.225	-0.127	-0.145
<i>V10</i>	<b>0.773</b>	0.001	-0.078	-0.070	-0.128
<i>V11</i>	<b>0.728</b>	-0.041	0.033	-0.017	-0.101
<i>V12</i>	<b>0.761</b>	-0.065	-0.076	-0.019	0.115
<i>V13</i>	-0.107	-0.027	-0.213	-0.182	<b>0.629</b>
<i>V14</i>	-0.149	-0.170	-0.081	-0.094	<b>0.665</b>
<i>V15</i>	-0.044	0.216	-0.060	-0.061	<b>0.344</b>
<i>V16</i>	-0.047	0.050	-0.060	-0.071	<b>0.646</b>
<i>V17</i>	-0.148	-0.146	<b>0.401</b>	<b>-0.466</b>	-0.102
<i>V18</i>	-0.146	-0.151	<b>0.546</b>	-0.323	-0.154
<i>V19</i>	-0.146	-0.107	<b>0.805</b>	-0.212	-0.088
<i>V20</i>	-0.053	0.005	<b>0.646</b>	-0.144	-0.093

La Tabla 28 muestra los indicadores de bondad de ajuste del modelo empleando el diseño intra-grupo en condición de respuesta honesta. En el presente análisis, los indicadores son semejantes o ligeramente mejores que los encontrados en la muestra total, aunque las diferencias no tienen relevancia práctica ya que se producen en las milésimas.

Tabla 28. Indicadores de Bondad de Ajuste del Modelo de Cinco Factores con Instrucciones de Respuesta Honesta al QI5F\_tri con un Diseño Intra-Grupo ( $N = 490$ )

---

Root Mean Square Error of Approximation (RMSEA) = 0.026
Estimated Non-Centrality Parameter (NCP) = 122.250
Degrees of Freedom = 100
Test of Approximate Fit $H_0$ : RMSEA < 0.05 ( $P = 1.000$ )
Normal Theory Chi Square with 100 degrees of freedom = 133.443 ( $P = 0.014317$ )
Chi-Square for independence model with 190 degrees of freedom = 4352.963
Non-Normed Fit Index (NNFI; Tucker & Lewis) = 0.985
Comparative Fit Index (CFI) = 0.992
Schwarz's Bayesian Information Criterion (BIC) = 876.771
Goodness of Fit Index (GFI) = 0.996
Adjusted Goodness of Fit Index (AGFI) = 0.992
Goodness of Fit Index without diagonal values (GFI) = 0.991
Adjusted Goodness of Fit Index without diagonal values (AGFI) = 0.983
T-Size Root Mean Square Error of Approximation (Ts-RMSEA) = 0.037
Maximum misspecification (T-size) = 0.138
T-Size Comparative Fit Index (Ts-CFI) = 0.981
Maximum misspecification (T-size) = 0.153
Root Mean Square of Residuals (RMSR) = 0.0211
Kelley's Expected mean value of RMSR for an acceptable model = 0.0452

---

En la Tabla 29 se presenta la matriz de cargas obtenida en condiciones de faking con el diseño intra-grupo. También en este caso se han obtenido resultados muy similares a los obtenidos en condiciones de faking con la muestra total. La estructura de cinco factores aparece claramente apoyada por los datos, aunque la magnitud de las cargas es ligeramente más pequeña que en condiciones honestas. También en el presente análisis dos de los compuestos de amigabilidad presentan cargas reducidas en el factor y negativas en el factor de conciencia.

Tabla 29. Matriz de Cargas Factoriales Rotadas del Análisis Factorial Exploratorio del Q15F\_tri Contestado Bajo Instrucciones de Faking, con un Diseño Intra-grupo ( $N = 490$ )

Variable	F1	F2	F3	F4	F5
<i>V1</i>	-0.074	-0.032	-0.031	-0.169	<b>0.489</b>
<i>V2</i>	0.109	0.047	0.120	0.070	<b>0.489</b>
<i>V3</i>	-0.015	-0.219	0.040	-0.087	<b>0.490</b>
<i>V4</i>	-0.190	0.261	0.013	-0.064	<b>0.418</b>
<i>V5</i>	<b>0.595</b>	-0.163	-0.145	-0.046	-0.094
<i>V6</i>	<b>0.619</b>	-0.189	-0.078	-0.099	0.056
<i>V7</i>	<b>0.431</b>	0.003	0.231	-0.000	0.036
<i>V8</i>	<b>0.650</b>	-0.119	0.031	-0.081	0.001
<i>V9</i>	-0.097	0.074	0.031	<b>0.593</b>	0.016
<i>V10</i>	-0.048	0.072	-0.060	<b>0.670</b>	0.054
<i>V11</i>	-0.010	0.171	-0.018	<b>0.692</b>	-0.009
<i>V12</i>	-0.071	-0.280	0.042	<b>0.629</b>	-0.191
<i>V13</i>	-0.149	<b>-0.671</b>	<b>0.034</b>	-0.264	-0.235
<i>V14</i>	-0.210	<b>-0.483</b>	<b>0.267</b>	-0.196	-0.226
<i>V15</i>	-0.080	0.091	<b>0.548</b>	0.014	0.099
<i>V16</i>	-0.137	-0.369	<b>0.281</b>	-0.156	-0.077
<i>V17</i>	-0.247	<b>0.489</b>	-0.036	-0.134	-0.032
<i>V18</i>	-0.215	<b>0.690</b>	0.034	-0.034	-0.038
<i>V19</i>	-0.244	<b>0.698</b>	0.080	-0.035	-0.152
<i>V20</i>	-0.187	<b>0.649</b>	0.024	-0.034	-0.050

Los indicadores de ajuste del modelo (Tabla 30) son, como en los casos anteriores, en conjunto buenos o excelentes, aunque ligeramente inferiores a los de la muestra total, pero debe tenerse en cuenta que el tamaño de la muestra es mucho más pequeño en el presente caso, con el consiguiente efecto del error de muestreo.

Tabla 30. Indicadores de Bondad de Ajuste del Modelo de Cinco Factores con Instrucciones de Respuesta que Inducían al QI5F\_tri, con un Diseño Intra-Grupo ( $N = 490$ )

---

Root Mean Square Error of Approximation (RMSEA) = 0.034
Estimated Non-Centrality Parameter (NCP) = 122.250
Degrees of Freedom = 100
Test of Approximate Fit $H_0$ : RMSEA < 0.05 ( $P = 0.996$ )
Normal Theory Chi Square with 100 degrees of freedom = 157.165 ( $P = 0.000235$ )
Chi-Square for independence model with 190 degrees of freedom = 3332.490
Non-Normed Fit Index (NNFI; Tucker & Lewis) = 0.965
Comparative Fit Index (CFI) = 0.982
Schwarz's Bayesian Information Criterion (BIC) = 900.494
Goodness of Fit Index (GFI) = 0.993
Adjusted Goodness of Fit Index (AGFI) = 0.987
Goodness of Fit Index without diagonal values (GFI) = 0.983
Adjusted Goodness of Fit Index without diagonal values (AGFI) = 0.968
T-Size Root Mean Square Error of Approximation (Ts-RMSEA) = 0.044
Maximum misspecification (T-size) = 0.195
T-Size Comparative Fit Index (Ts-CFI) = 0.965
Maximum misspecification (T-size) = 0.211
Root Mean Square of Residuals (RMSR) = 0.0246
Kelley's Expected mean value of RMSR for an acceptable model = 0.0452

---

### *1.3 Análisis factorial exploratorio en el diseño inter-grupos*

Finalmente, los dos últimos análisis EFA se efectuaron con un diseño entre-grupos, lo que permite comprobar el potencial efecto del diseño de investigación sobre el ajuste del modelo, al comparar los resultados de estos análisis con los del diseño intra-grupo. En la Tabla 31 se presentan los estadísticos descriptivos de cada una de las variables para esta muestra. Se pueden consultar las matrices de correlaciones factorizadas del diseño intra-grupo en el apartado Anexos 6 y 7.

Tabla 31. Estadísticos Descriptivos de la Variables en el Diseño Inter-Grupos

Variable	Honesta (N = 449)		Faking (N = 202)	
	Media	SD	Media	SD
V1	5.93	5.75	6.03	4.08
V2	5.78	5.90	7.17	5.09
V3	5.05	5.57	5.50	3.90
V4	6.02	5.45	7.00	4.01
V5	7.32	7.82	5.97	4.77
V6	7.29	7.07	7.07	4.01
V7	6.49	5.60	6.84	4.41
V8	7.53	6.35	7.05	4.51
V9	6.64	8.06	7.96	7.56
V10	7.18	7.02	7.48	4.56
V11	6.44	10.10	7.66	6.36
V12	6.55	8.85	6.67	7.56
V13	8.32	7.00	5.76	5.07
V14	7.45	6.11	6.25	5.00
V15	6.36	4.19	7.46	4.18
V16	7.53	5.94	6.75	4.47
V17	6.04	7.83	6.58	5.13
V18	6.52	5.50	7.31	4.71
V19	7.45	8.84	9.05	6.02
V20	6.30	5.92	8.10	6.09

Los estadísticos descriptivos muestran que, de nuevo, las medias en la condición faking son más altas que en la condición honesta, con valores promedio de 6.98 y 6.71, respectivamente, mientras que, para la desviación típica, los valores tienden a ser más pequeños para la condición faking, obteniéndose una SD promedio de 5.07 para esta condición y de 6.74 para la condición honesta. Estos resultados de nuevo son semejantes a los obtenidos previamente para las otras muestras.

Las Tablas 32 y 33 muestran los resultados en condiciones de respuesta honesta y con diseño entre-grupos. La primera de ellas incluye la matriz de cargas factoriales y la segunda presenta los estadísticos de bondad de ajuste.

Puede comprobarse en la Tabla 32 que todas las variables cargan en su hipotético factor con cargas relevantes y positivas y que no presentan cargas secundarias positivas en los restantes factores. También puede observarse que las restantes cargas no son significativas y la gran mayoría de ellas negativas. Por tanto, los resultados, una vez más, indican que una estructura factorial de cinco factores ortogonales se ajusta a los datos.

Tabla 32. Matriz de Cargas Factoriales Rotadas del Análisis Factorial Exploratorio del Q15F\_tri Contestado Bajo Instrucciones Honestas, con un Diseño Entre-Grupos ( $N = 449$ )

Variable	F1	F2	F3	F4	F5
<i>V1</i>	<b>0.578</b>	-0.182	-0.023	-0.022	-0.096
<i>V2</i>	<b>0.577</b>	-0.067	-0.165	0.204	0.003
<i>V3</i>	<b>0.556</b>	0.089	-0.231	-0.052	-0.048
<i>V4</i>	<b>0.625</b>	0.017	0.016	-0.184	-0.049
<i>V5</i>	-0.057	-0.134	-0.139	<b>0.732</b>	-0.048
<i>V6</i>	-0.011	-0.117	-0.232	<b>0.756</b>	-0.033
<i>V7</i>	-0.083	0.070	-0.191	<b>0.495</b>	-0.031
<i>V8</i>	0.023	-0.094	-0.115	<b>0.728</b>	-0.127
<i>V9</i>	-0.014	-0.104	-0.142	-0.077	<b>0.707</b>
<i>V10</i>	0.021	-0.110	-0.005	-0.092	<b>0.751</b>
<i>V11</i>	-0.045	-0.098	0.029	0.016	<b>0.806</b>
<i>V12</i>	-0.048	0.086	-0.119	-0.047	<b>0.788</b>
<i>V13</i>	0.014	<b>0.729</b>	-0.169	-0.171	-0.111
<i>V14</i>	-0.199	<b>0.674</b>	0.037	-0.107	-0.062
<i>V15</i>	0.288	<b>0.442</b>	-0.000	-0.117	0.026
<i>V16</i>	0.044	<b>0.612</b>	-0.043	-0.039	-0.097
<i>V17</i>	-0.059	-0.047	<b>0.457</b>	-0.291	-0.084
<i>V18</i>	-0.087	-0.058	<b>0.577</b>	-0.282	-0.156
<i>V19</i>	-0.086	-0.087	<b>0.810</b>	-0.174	-0.080
<i>V20</i>	-0.008	-0.002	<b>0.602</b>	-0.206	-0.030

Por lo que se refiere a los estadísticos de bondad de ajuste, se presentan en la Tabla 33. Al igual que en todos los casos precedentes, son buenos o excelentes, con magnitudes semejantes a los de los análisis previos.

Tabla 33. Indicadores de Bondad de Ajuste del Modelo de Cinco Factores con Instrucciones de Respuesta que Inducía Honestidad al QI5F\_tri, con un Diseño Entre-Grupos ( $N = 449$ )

---

Root Mean Square Error of Approximation (RMSEA) = 0.035
Estimated Non-Centrality Parameter (NCP) = 112.000
Degrees of Freedom = 100
Test of Approximate Fit $H_0$ : RMSEA < 0.05 ( $P = 0.993$ )
Normal Theory Chi Square with 100 degrees of freedom = 153.727 ( $P = 0.000455$ )
Chi-Square for independence model with 190 degrees of freedom = 4028.950
Non-Normed Fit Index (NNFI; Tucker & Lewis) = 0.973
Comparative Fit Index (CFI) = 0.986
Schwarz's Bayesian Information Criterion (BIC) = 886.569
Goodness of Fit Index (GFI) = 0.995
Adjusted Goodness of Fit Index (AGFI) = 0.990
Goodness of Fit Index without diagonal values (GFI) = 0.989
Adjusted Goodness of Fit Index without diagonal values (AGFI) = 0.979
T-Size Root Mean Square Error of Approximation (Ts-RMSEA) = 0.045
Maximum misspecification (T-size) = 0.204
T-Size Comparative Fit Index (Ts-CFI) = 0.972
Maximum misspecification (T-size) = 0.221
Root Mean Square of Residuals (RMSR) = 0.0222
Expected mean value of RMSR for an acceptable model = 0.0472

---

Por último, los resultados de EFA con las respuestas bajo instrucciones de faking se presentan en las Tablas 34 y 35. En la primera de ellas se muestra la matriz de cargas factoriales que, como puede apreciarse, vuelven a poner de manifiesto que una estructura de cinco factores ortogonales se ajusta a los datos. No obstante, hay que señalar que en el presente análisis tres de las variables (V1, V13 y V14) muestra una carga factorial no significativa en su correspondiente factor, aunque no cargan de modo significativo en ningún otro factor.



Tabla 34. Matriz de Cargas Factoriales Rotadas del Análisis Factorial Exploratorio del Q15F\_tri Contestado Bajo Instrucciones de Faking, con un Diseño Entre-Grupos ( $N = 202$ )

Variable	F1	F2	F3	F4	F5
<i>V1</i>	<b>0.153</b>	-0.093	-0.030	-0.156	-0.115
<i>V2</i>	<b>0.598</b>	0.058	0.128	-0.045	0.107
<i>V3</i>	<b>0.352</b>	-0.244	-0.250	-0.147	-0.173
<i>V4</i>	<b>0.659</b>	0.140	-0.127	0.144	0.019
<i>V5</i>	-0.028	-0.267	<b>0.331</b>	-0.157	-0.191
<i>V6</i>	-0.112	-0.399	<b>0.482</b>	-0.075	-0.038
<i>V7</i>	0.052	-0.115	<b>0.584</b>	-0.116	0.064
<i>V8</i>	-0.043	-0.209	<b>0.333</b>	-0.216	-0.048
<i>V9</i>	0.006	0.161	0.021	<b>0.674</b>	0.042
<i>V10</i>	0.019	0.004	-0.039	<b>0.563</b>	-0.144
<i>V11</i>	0.111	0.110	0.012	<b>0.645</b>	-0.049
<i>V12</i>	0.020	-0.197	-0.274	<b>0.617</b>	0.047
<i>V13</i>	<b>-0.319</b>	<b>-0.320</b>	-0.238	-0.310	<b>0.141</b>
<i>V14</i>	<b>-0.337</b>	-0.141	-0.254	-0.196	<b>0.141</b>
<i>V15</i>	0.189	0.047	0.063	-0.017	<b>0.576</b>
<i>V16</i>	-0.225	-0.147	-0.113	-0.084	<b>0.462</b>
<i>V17</i>	-0.055	<b>0.428</b>	-0.050	0.034	-0.144
<i>V18</i>	0.114	<b>0.524</b>	-0.192	-0.019	-0.079
<i>V19</i>	0.099	<b>0.647</b>	-0.009	0.040	0.037
<i>V20</i>	0.087	<b>0.585</b>	-0.141	0.075	0.011

Por lo que se refiere a los estadísticos de bondad de ajuste (Tabla 35), son todos ellos buenos o excelentes, destacando particularmente la RMSEA cuyo valor indica un mejor ajuste que en el caso de la muestra en condiciones honestas (.023 vs. .035). Por tanto, en conjunto hay que señalar que este último análisis también proporciona apoyo empírico a la estructura de cinco factores y a la equivalencia (invarianza) de dicha estructura en condiciones honestas y de faking.

Tabla 35. Indicadores de Bondad de Ajuste del Modelo de Cinco Factores con Instrucciones de Respuesta que Inducían Faking al QI5F\_tri, con un Diseño Entre-Grupos ( $N = 202$ )

---

Root Mean Square Error of Approximation (RMSEA) = 0.023
Estimated Non-Centrality Parameter (NCP) = 50.250
Degrees of Freedom = 100
Test of Approximate Fit $H_0$ : RMSEA < 0.05; $P = 0.982$
Normal Theory Chi Square with 100 degrees of freedom = 110.653 ( $P = 0.219177$ )
Chi-Square for independence model with 190 degrees of freedom = 983.161
Non-Normed Fit Index (NNFI; Tucker & Lewis) = 0.974
Comparative Fit Index (CFI) = 0.987
Schwarz's Bayesian Information Criterion (BIC) = 747.645
Goodness of Fit Index (GFI) = 0.984
Adjusted Goodness of Fit Index (AGFI) = 0.970
Goodness of Fit Index without diagonal values (GFI) = 0.952
Adjusted Goodness of Fit Index without diagonal values (AGFI) = 0.908
T-Size Root Mean Square Error of Approximation (Ts-RMSEA) = 0.045
Maximum misspecification (T-size) = 0.20
T-Size Comparative Fit Index (Ts-CFI) = 0.931
Maximum misspecification (T-size) = 0.233
Root Mean Square of Residuals (RMSR) = 0.0354
Kelley's Expected mean value of RMSR for an acceptable model = 0.0705

---

#### 1.4. Discusión de los Resultados de los EFA

En resumen, los seis análisis EFA realizados, comparando la estructura del QI5F\_tri en condiciones de respuesta honesta y respuesta con instrucciones de faking, empleando un diseño entre-grupos y un diseño intra-grupo, produjeron resultados muy semejantes, que podemos establecer en los siguientes puntos:

1. La estructura de cinco factores del QI5F\_tri se obtiene en todos los análisis, con independencia de la condición de respuesta (honesto o faking) o el tipo de diseño (entre-grupos o intra-grupo).
2. Los estadísticos de bondad de ajuste del modelo, tanto en condición honesta como faking, y tanto en el diseño entre-grupos, como en el diseño intra-grupo, son buenos o excelentes, de acuerdo a los puntos de corte establecidos por Hu y Blenter (1999) y de acuerdo a los niveles de significación establecidos con los estadísticos T de Chan et al. (2016). En condiciones de faking no se aprecia que la magnitud de los estadísticos se haya visto afectada de modo relevante, dando

el caso de que algunos indicadores eran mejores en condiciones de faking que honestas.

3. El tamaño de las cargas factoriales es algo mayor en la condición honesta que en la condición de faking, aunque solo se apreció un cambio relevante en dos de los clusters del factor de amigabilidad.

4. La comparación de las cargas factoriales y de los estadísticos de ajuste de las tres muestras independientes (honestas entre grupos, faking entre grupos, e muestra intra-grupo) apoya la invarianza de la medida en el caso del QI5F\_tri, apoyando la hipótesis de que el faking no produce alteraciones significativas en el caso de un cuestionario de EF quasi-ipsativo con independencia de las medidas.

5. No puede generalizarse los resultados a otros tipos de cuestionarios de EF, como los ipsativos o quasi-ipsativos con dependencia, aunque se puede especular que el formato de EF es resistente a las modificaciones de la estructura que el faking produce en los cuestionarios SS.

6. Debe tenerse en cuenta, además, que los análisis efectuados se realizaron en condiciones que favorecen la discrepancia en las estructuras factoriales, ya que los clusters de ítems fueron creados mediante una asignación arbitraria. La utilización de los compuestos propias de las Big Five medidas por el QI5F\_tri, sin duda, hubiera contribuido a una mayor equivalencia (invarianza) de la medida y a mejores estadísticos de ajuste, ya que los errores de medida y la varianza de factores específicos se hubieran reducido. No se emplearon las para establecer un potencial mayor desacuerdo entre las condiciones honestas y faking.

## 2. ANÁLISIS DE COMPONENTES PRINCIPALES DE LA ESTRUCTURA DEL QI5F\_TRI

Algunos autores como Costa y McCrae (1992) y Goldberg (1994) sostienen que el análisis de componentes principales (ACP) puede ser más adecuado que el AFE para establecer la estructura de los cuestionarios de personalidad. Por ello, y como un complemento a los análisis presentados en la sección previa, seguidamente se presentan los resultados de los análisis de componentes principales realizados con las mismas matrices de datos de dichos AFE, aunque no se esperan cambios sustanciales en los resultados, más allá de los debidos a la naturaleza las matrices de correlación empleadas en el AFE y el ACP.

Como en el caso de los AFE, en todos los casos en los análisis de componentes principales se extrajeron con cinco componentes y se empleó la rotación Varimax. Al igual que los AFE, los ACP se realizaron con el programa FACTOR (Lorenzo-Seva y Ferrando, 2018).

### 2.1 Análisis de componentes principales en la muestra total

En la Tabla 36 se presentan las cargas factoriales de las variables de personalidad en los cinco componentes para la muestra total del estudio en la condición honesta.

Tabla 36. Matriz de Cargas Rotadas de Cinco Componentes Principales en la Condición Honesta para la Muestra Total ( $N = 939$ )

Variable	C1	C2	C3	C4	C5
<i>V1</i>	<b>0.650</b>	-0.093	-0.158	-0.113	-0.213
<i>V2</i>	<b>0.700</b>	-0.116	0.001	0.270	-0.000
<i>V3</i>	<b>0.605</b>	-0.356	-0.093	-0.154	0.003
<i>V4</i>	<b>0.701</b>	0.036	-0.082	-0.226	0.057
<i>V5</i>	-0.094	-0.139	-0.118	<b>0.765</b>	-0.170
<i>V6</i>	-0.013	-0.279	-0.075	<b>0.717</b>	-0.174
<i>V7</i>	-0.1.1	-0.138	-0.020	<b>0.644</b>	0.075
<i>V8</i>	0.009	-0.115	-0.153	<b>0.772</b>	-0.094
<i>V9</i>	-0.045	-0.196	<b>0.771</b>	--0.115	-0.135
<i>V10</i>	0.010	-0.026	<b>0.820</b>	-0.081	-0.106
<i>V11</i>	-0.046	0.058	<b>0.827</b>	0.009	-0.082
<i>V12</i>	-0.093	-0.132	<b>0.811</b>	-0.072	0.082
<i>V13</i>	-0.110	-0.332	-0.173	-0.273	<b>0.647</b>
<i>V14</i>	-0.291	-0.109	-0.162	-0.158	<b>0.683</b>
<i>V15</i>	0.365	0.165	0.075	0.047	<b>0.684</b>
<i>V16</i>	-0.028	-0.105	-0.106	-0.080	<b>0.728</b>
<i>V17</i>	-0.134	<b>0.443</b>	-0.156	<b>-0.461</b>	-0.134
<i>V18</i>	-0.143	<b>0.612</b>	-0.185	-0.346	-0.142
<i>V19</i>	-0.133	<b>0.781</b>	-0.151	-0.234	-0.103
<i>V20</i>	0.004	<b>0.768</b>	-0.032	-0.152	-0.013

Tal como esperábamos los resultados replican la distribución de factores descrita en la Tabla 20. Los cuatro compuestos creados para el factor estabilidad emocional, cargan de modo significativo en el componente 1. En el componente 2 cargan exclusivamente las variables del factor conciencia. El componente 3 del modelo incluye las cargas factoriales del factor apertura a la experiencia. Los clusters del factor extraversión, cargan en el componente 4 y, finalmente, los clusters que representan al

factor amigabilidad cargan juntos en el componente 5. También puede observarse que en todos los casos las variables muestran signos positivos de correlacionan con el factor (cargas positivas).

Por tanto, la matriz de correlaciones muestra una estructura de cinco componentes principales en la condición honesta. Por otra parte, apenas hay cargas relevantes en los componentes, además de las propias de cada factor de personalidad. Por ejemplo, en el componente 1, las restantes cargas oscilan entre  $-.143$  y  $-.365$ . En el componente 2 van de  $-.332$  a  $-.058$ . En el componente 3 van desde  $-.185$  a  $-.075$ . En el componente 4 van desde  $-.461$  a  $.270$  y, finalmente, en el componente 5 las cargas oscilan entre  $-.213$  y  $.082$ . Por tanto, puede afirmarse que la estructura de cinco componentes explica adecuadamente la varianza de las 20 variables.

Respecto a la condición faking, en la Tabla 37 se presenta la matriz de cargas rotadas de los cinco componentes principales obtenida para esta condición. Como puede observarse los resultados replican la estructura de cinco factores obtenida en la condición honesta. En el componente primero cargan juntos los clusters que representan al factor apertura a la experiencia. En el segundo componente cargan los compuestos del factor extraversión. En el componente tercero cargan juntos los clusters del factor amigabilidad y, por último, en los componentes cuarto y quinto, cargan de modo significativo y único los compuestos que representan a los factores conciencia y estabilidad emocional, respectivamente.

Al igual que en la condición honesta, apenas hay cargas relevantes en los componentes, además de las propias de cada factor de personalidad. Por ejemplo, en el componente 1, las restantes cargas oscilan entre  $-.297$  y  $.080$ . En el componente 2 van de  $-.280$  a  $.172$ . En el componente 3 van desde  $-.234$  a  $.343$ . En el componente 4 van desde  $-.402$  a  $.286$  y, finalmente, en el componente 5 las cargas oscilan entre  $-.257$  y  $.092$ . Por tanto, también en el caso de la condición faking puede afirmarse que la estructura de cinco componentes explica adecuadamente la varianza de las 20 variables.

Tabla 37. Matriz de Cargas Rotadas de Cinco Componentes Principales en la Condición Faking para la Muestra Total ( $N = 692$ )

Variable	C1	C2	C3	C4	C5
<i>V1</i>	-0.200	-0.086	-0.100	-0.070	<b>0.622</b>
<i>V2</i>	0.080	0.172	0.259	0.076	<b>0.629</b>
<i>V3</i>	-0.094	-0.031	-0.028	-0.269	<b>0.634</b>
<i>V4</i>	0.020	-0.196	0.075	0.286	<b>0.583</b>
<i>V5</i>	-0.079	<b>0.637</b>	-0.234	-0.171	-0.100
<i>V6</i>	-0.104	<b>0.675</b>	-0.116	-0.229	0.034
<i>V7</i>	-0.032	<b>0.636</b>	0.343	0.034	-0.004
<i>V8</i>	-0.127	<b>0.686</b>	-0.004	-0.127	-0.028
<i>V9</i>	<b>0.716</b>	-0.081	0.067	0.095	-0.028
<i>V10</i>	<b>0.740</b>	-0.046	-0.096	0.046	0.052
<i>V11</i>	<b>0.754</b>	-0.007	-0.009	0.170	-0.019
<i>V12</i>	<b>0.698</b>	-0.146	0.024	-0.298	-0.151
<i>V13</i>	-0.297	-0.214	<b>0.010</b>	<b>-0.667</b>	-0.228
<i>V14</i>	-0.237	-0.280	<b>0.245</b>	<b>-0.501</b>	-0.257
<i>V15</i>	0.015	-0.010	<b>0.826</b>	0.084	0.092
<i>V16</i>	-0.161	-0.185	<b>0.470</b>	-0.402	-0.186
<i>V17</i>	-0.147	-0.271	-0.071	<b>0.550</b>	-0.077
<i>V18</i>	-0.038	-0.233	0.043	<b>0.715</b>	-0.048
<i>V19</i>	-0.033	-0.223	0.104	<b>0.730</b>	-0.123
<i>V20</i>	-0.018	-0.222	0.060	<b>0.679</b>	-0.053

Si se comparan los resultados del ACP en la condición honesta y la condición faking, puede observarse que, tal y como esperábamos, la estructura se mantiene en ambas condiciones en el cuestionario de EF, aunque el orden de componentes resultante no es el mismo. Los componentes primero y segundo de la condición honesta son el quinto y cuarto de la condición faking y los restantes componentes de la condición honesta, tercero, cuarto y quinto componente, resultan ser el primero, segundo y tercero, respectivamente, en la condición faking. En todo caso, la estructura de componentes principales resultante coincide totalmente en ambas condiciones cuando se evalúa la personalidad con cuestionarios de EF.

Tabla 38. Coeficientes de Ajuste del Modelo del Análisis de Componentes Principales

Condición Experimental	RMSR
<i>Honesta</i>	0.0565
<i>Faking</i>	0.0654

*Nota.* RMSR = Raíz Cuadrada Media de los Residuos Estandarizados

También se han examinado los índices de ajuste del modelo obtenidos en este ACP para las dos condiciones experimentales. Los resultados se muestran en la Tabla 38. Como se puede apreciar, los valores de la raíz cuadrada media de los residuos estandarizados (RMSR) sugieren un ajuste aceptable del modelo a los datos en ambos casos, aunque ligeramente mejor en el caso de la muestra honesta. No obstante, hay que señalar que la muestra honesta y la muestra en condición faking no son exactamente equivalentes, ya que tanto el tamaño como las características de los participantes varían, lo que también puede contribuir a explicar la diferencia en la RMSR.

## 2.2 Análisis de componentes principales en el diseño intra-grupo

Al igual que se hizo para la muestra total, este ACP se ha realizado con cinco componentes. La matriz de las cargas factoriales rotadas para el diseño intra-grupo en la condición honesta se presenta en la Tabla 39.

Los resultados muestran que la estructura factorial obtenida replica la encontrada en la muestra total para la condición honesta. No obstante, el orden de los factores resultantes varía, a excepción de los componentes tercero y cuarto. Así, en el componente 1 cargan juntos los clusters del factor amigabilidad. Los compuestos del factor estabilidad emocional, cargan de modo único y significativo en el componente 2. Por su parte, el componente 3 incluye las cargas factoriales de los compuestos de apertura a la experiencia y en el componente 4 cargan juntos los compuestos del factor extraversión. Por último, los clusters de conciencia cargan de modo relevante en el componente 5. En conclusión, la estructura factorial se mantiene también en esta condición experimental cuando se usan cuestionarios de personalidad de EF.

Asimismo, apenas hay cargas relevantes en los componentes, a excepción de las propias de cada factor de personalidad. En el componente 1, las restantes cargas oscilan entre -.220 y .108. En el componente 2 van de -.277 a .349. En el componente 3 van desde



-.206 a .056. En el componente 4 van desde -.549 a .262 y, finalmente, en el componente 5 las cargas oscilan entre -.376 y .189. Nuevamente, se puede afirmar que la estructura de cinco componentes explica adecuadamente la varianza de las 20 variables.

Tabla 39. Matriz de Cargas Rotadas de Cinco Componentes Principales en la Condición Honesta para la Muestra Intra-Grupo ( $N = 490$ )

Variable	C1	C2	C3	C4	C5
<i>V1</i>	-0.220	<b>0.623</b>	-0.158	-0.163	-0.150
<i>V2</i>	-0.015	<b>0.696</b>	-0.029	0.262	-0.128
<i>V3</i>	-0.032	<b>0.615</b>	-0.081	-0.130	-0.338
<i>V4</i>	0.103	<b>0.697</b>	-0.103	-0.207	0.082
<i>V5</i>	-0.162	-0.128	-0.139	<b>0.767</b>	-0.120
<i>V6</i>	-0.171	0.006	-0.092	<b>0.703</b>	-0.260
<i>V7</i>	0.052	-0.095	0.005	<b>0.658</b>	-0.091
<i>V8</i>	-0.081	-0.005	-0.137	<b>0.759</b>	-0.113
<i>V9</i>	-0.139	-0.059	<b>0.765</b>	-0.108	-0.211
<i>V10</i>	-0.109	0.004	<b>0.830</b>	-0.061	-0.052
<i>V11</i>	-0.085	-0.048	<b>0.799</b>	-0.011	0.063
<i>V12</i>	0.108	-0.107	<b>0.799</b>	-0.044	-0.112
<i>V13</i>	<b>0.609</b>	-0.122	-0.167	-0.269	-0.376
<i>V14</i>	<b>0.662</b>	-0.277	-0.206	-0.156	-0.192
<i>V15</i>	<b>0.682</b>	0.349	0.056	0.098	0.189
<i>V16</i>	<b>0.727</b>	-0.039	-0.077	-0.106	-0.130
<i>V17</i>	-0.160	-0.164	-0.183	-0.549	<b>0.383</b>
<i>V18</i>	-0.173	-0.162	-0.171	-0.369	<b>0.591</b>
<i>V19</i>	-0.100	-0.015	-0.182	-0.284	<b>0.757</b>
<i>V20</i>	-0.045	-0.002	-0.056	-0.146	<b>0.780</b>

Los resultados obtenidos para la condición faking siguen la línea de los descritos en los análisis precedentes. La matriz de cargas rotadas para esta condición se presenta en la Tabla 40.

Tabla 40. Matriz de Cargas Rotadas de Cinco Componentes Principales en la Condición Faking para la Muestra Intra-Grupo ( $N = 490$ )

Variable	C1	C2	C3	C4	C5
<i>V1</i>	-0.092	-0.029	<b>0.669</b>	-0.084	-0.182
<i>V2</i>	0.148	0.059	<b>0.623</b>	0.201	0.118
<i>V3</i>	0.029	-0.243	<b>0.655</b>	0.052	-0.068
<i>V4</i>	-0.215	0.302	<b>0.556</b>	0.026	-0.053
<i>V5</i>	<b>0.685</b>	-0.145	-0.112	-0.194	-0.062
<i>V6</i>	<b>0.707</b>	-0.172	0.066	-0.116	-0.110
<i>V7</i>	<b>0.589</b>	0.043	-0.012	<b>0.444</b>	0.006
<i>V8</i>	<b>0.734</b>	-0.093	-0.013	0.011	-0.094
<i>V9</i>	-0.098	0.072	-0.006	0.067	<b>0.713</b>
<i>V10</i>	-0.057	0.071	0.055	-0.076	<b>0.758</b>
<i>V11</i>	-0.005	0.181	-0.042	0.007	<b>0.758</b>
<i>V12</i>	-0.089	-0.324	-0.200	0.014	<b>0.690</b>
<i>V13</i>	-0.158	-0.721	-0.192	-0.013	-0.280
<i>V14</i>	-0.255	-0.574	-0.216	<b>0.271</b>	-0.227
<i>V15</i>	-0.068	0.085	0.078	<b>0.819</b>	0.024
<i>V16</i>	-0.178	-0.465	-0.070	<b>0.410</b>	-0.189
<i>V17</i>	-0.297	<b>0.546</b>	-0.064	-0.012	-0.191
<i>V18</i>	-0.229	<b>0.737</b>	-0.067	0.096	-0.041
<i>V19</i>	-0.274	<b>0.716</b>	-0.167	0.110	-0.055
<i>V20</i>	-0.217	<b>0.694</b>	-0.072	0.055	-0.053

Como se puede apreciar de nuevo se replica estructura factorial propuesta de cinco componentes, aunque se producen diferencias en el orden de cargas respecto a la condición honesta y a los resultados de la muestra total. En este caso, los clusters del factor extraversión cargan juntos en el primer componente. En el segundo componente cargan los compuestos del factor conciencia. En el componente 3 cargan juntos los clusters de estabilidad emocional. En el cuarto componente cargan juntos los compuestos del factor amigabilidad. Finalmente, en el componente 5 cargan juntos los compuestos de apertura a la experiencia. Como se puede apreciar, en esta condición, la V13 no carga significativamente en ningún factor, sus cargas son negativas en todos los componentes. No obstante, de nuevo, se obtiene una estructura factorial de cinco componentes tal y como se esperaba. Estos resultados muestran que cuando se usan cuestionarios de EF se produce invarianza en la estructura factorial en presencia de faking, incluso en el diseño intra-grupo. En consecuencia, la estructura factorial coincide en todas las condiciones y muestras analizadas.

Asimismo, se han examinado los índices de ajuste del modelo obtenidos en el análisis factorial exploratorio para ambas condiciones. Los valores sugieren un buen

ajuste del modelo a los datos, con valores de RMSR próximos a .06 en ambos casos, como se puede observar en la Tabla 41.

Tabla 41. Coeficientes de Ajuste del Modelo del Análisis Factorial Exploratorio de la Muestra Intra-Grupo

Condición experimental	RMSR
<i>Honesta</i>	0.0582
<i>Faking</i>	0.0649

*Nota.* RMSR = media de la raíz de las diferencias.

### 2.3 Análisis de la congruencia de estructura en los análisis de componentes principales

Se ha utilizado el coeficiente de congruencia de Burt-Tucker ( $r_c$ ) para el cálculo de la congruencia factorial, siguiendo la recomendación de Cattell (1971). Este coeficiente permite conocer la similitud del patrón de cargas, de tal modo que si los coeficientes que se obtengan entre factores convergentes son altos (en torno a .90) indican que los factores emparejados son similares, especialmente si los coeficientes entre factores divergentes son bajos (inferiores a .40).

Los coeficientes de Burt-Tucker se han calculado para cada par de factores convergentes en cada una de las condiciones experimentales, honesta y faking, y entre las condiciones experimentales para obtener todo el conjunto de coeficientes posible. Para su cálculo se han usado los pesos factoriales de las matrices de las cargas rotadas obtenidos en los análisis factoriales de la muestra total. En las Tablas 42 y 43 se presentan los coeficientes de Burt-Tucker calculados para las condiciones honesta y faking, respectivamente.

Tabla 42. Coeficientes de Congruencia entre los Factores de Personalidad de la Condición Honesta ( $N = 939$ )

	EE	EX	AE	A	C
EE	-				
EX	-.01	-			
AE	-.07	-.06	-		
A	-.04	-.17	-.12	-	
C	-.18	-.40	-.14	-.18	-

*Nota.* EE = estabilidad emocional; EX = extraversión; AE = apertura a la experiencia; A = amigabilidad; C = conciencia.

Tabla 43. Coeficientes de Congruencia entre los Factore de Personalidad de la Condición Faking ( $N = 692$ )

	EE	EX	AE	A	C
EE	-				
EX	.00	-			
AE	-.02	-.07	-		
A	.03	-.10	-.04	-	
C	.09	-.23	.13	-.01	-

*Nota.* EE = estabilidad emocional; EX = extraversión; AE = apertura a la experiencia; A = amigabilidad; C = conciencia.

Como se puede apreciar en las tablas, los coeficientes entre factores convergentes son muy altos en comparación al resto de coeficientes, por lo tanto, la estructura factorial se mantiene en los cuestionarios de EF a pesar de los efectos del faking.

No obstante, es importante analizar si la congruencia factorial se mantiene entre factores convergentes, pero de condiciones diferentes. En la Tabla 44 se recogen los coeficientes de congruencia factorial entre ambas condiciones.

Tabla 44. Coeficientes de Congruencia entre los Factores de Personalidad de la Condiciones Honesta y Faking

	EE	EX	AE	A	C
EE	<b>.96</b>	.01	-.07	-.24	-.18
EX	.00	<b>.96</b>	-.08	-.28	-.40
AE	-.04	-.06	<b>.98</b>	-.24	-.04
A	.21	-.02	-.01	<b>.73</b>	.10
C	.03	-.28	-.01	-.48	<b>.89</b>
Media	<b>.90</b>				
M. Absoluta	.14				
M. Relativa	-.10				

*Nota.* EE = estabilidad emocional; EX = extraversión; AE = apertura a la experiencia; A = amigabilidad; C = conciencia.

Los coeficientes obtenidos muestran valores mucho más altos entre factores convergentes de condiciones experimentales diferentes que entre los factores divergentes. En promedio el coeficiente de congruencia de los componentes similares es de .90, mientras que el promedio de los factores divergentes es de .14 en términos absolutos y -.10 en términos relativos.

En resumen, los coeficientes de congruencia obtenidos apoyan que la estructura factorial de la personalidad es robusta en los cuestionarios de EF en todos los casos, independientemente de la condición o el diseño experimental.

### 3. ANÁLISIS FACTORIAL CONFIRMATORIO DE LA ESTRUCTURA DEL QI5F\_TRI

También se ha realizado un análisis factorial confirmatorio de componentes. El análisis se hizo con el objetivo de corroborar la estructura factorial propuesta de cinco factores (Tabla 20) en las dos condiciones experimentales analizadas (honesta y faking).

#### 3.1 Análisis factorial confirmatorio en la muestra total

Primero se presentan los coeficientes de ajuste obtenidos en el análisis factorial confirmatorio para la muestra total de este estudio (Tabla 45).

Tabla 45. Coeficientes de Ajuste del Modelo del Análisis Factorial Confirmatorio para la Muestra Total

Condición	$\chi^2$	DF	X <sup>2</sup> /DF	RMSEA	RMR	GFI	CFI	NNFI
<i>Honesta</i>	714.42	160	4.47	.061	.063	.93	.90	.88
<i>Faking</i>	529.57	160	3.31	.058	.068	.93	.86	.83

*Nota.*  $\chi^2$  = prueba chi cuadrado; DF = grados de libertad; RMSEA = raíz del residuo cuadrático promedio de aproximación; RMR = raíz del residuo cuadrático promedio; GFI = índice de bondad; CFI = índice de bondad de ajuste comparativo; NNFI = índice de ajuste no normalizado.

En el presente caso, para la interpretación de los resultados son particularmente relevantes los valores de índices de ajuste basados en los residuos, es decir RMSEA y SRMR. Como puede apreciarse, ambos valores estimadores de ajuste presentan valores en torno a .06, lo que de acuerdo con los puntos de corte sugeridos por Hu y Bentler (1999) indicaría un buen ajuste del modelo a los datos. Los índices de ajuste relativos NNFI, CFI, y GFI valores próximos a .90, lo que en combinación con los valores de RMSEA y SRMR refuerza la interpretación de un buen ajuste a los datos. Por ejemplo, CFI = .90 para la condición honesta y CFI = .86 en la condición faking o GFI, cuyo valor es .93 en ambos casos. Estos resultados indican un buen ajuste del modelo a los datos. Por tanto, podemos concluir que el modelo presenta un excelente ajuste a los datos tanto en la condición honesta como en la condición faking cuando se usan medidas de EF. Por tanto, estos resultados apoyan la *H1*.

### 3.2 Análisis factorial confirmatorio en la muestra intra-grupo

Para la muestra intra-grupo también se ha realizado un análisis factorial confirmatorio de componentes para corroborar la estructura factorial que se ha propuesto de cinco componentes (Tabla 20). Los índices de ajuste del modelo obtenidos se resumen en la Tabla 46.

Tabla 46. Coeficientes de Ajuste del Modelo del Análisis Factorial Confirmatorio para la Muestra Intra-Grupo

Condición	$\chi^2$	DF	X <sup>2</sup> /DF	RMSEA	RMR	GFI	CFI	NNFI
<i>Honesta</i>	419.20	160	2.62	.058	.054	.92	.91	.89
<i>Faking</i>	451.24	160	2.82	.062	.075	.91	.86	.83

*Nota.*  $\chi^2$  = prueba chi cuadrado; DF = grados de libertad; RMSEA = raíz del residuo cuadrático promedio de aproximación; RMR = raíz del residuo cuadrático promedio; GFI = índice de bondad; CFI = índice de bondad de ajuste comparativo; NNFI = índice de ajuste no normalizado.

Como se puede apreciar, los índices de ajuste basados en los residuos, RMSEA y RMR son muy aceptables, con valores inferiores a .075. Asimismo, todos los índices de bondad de ajuste muestran valores en torno a .90. En base a estos resultados podemos concluir que el modelo presenta un excelente ajuste a los datos tanto en la condición honesta como en la condición *faking* para el diseño intra-grupo. Estos resultados, nuevamente, apoyan la *H2*.

## DISCUSIÓN

El objetivo principal de este estudio ha sido examinar a invarianza de la medida de los Big Five, es decir, comprobar si las medidas en diferentes grupos son comparables entre sí. El modo más frecuente de examinar la invarianza de la medida es realizar una serie de análisis estructurales (p. ej., análisis factorial exploratorio, análisis de componentes principales, análisis factorial confirmatorio), obtener la matriz de cargas factoriales y estimar diversos estadísticos de ajuste del modelo a los datos. Hasta hace pocos años, no se disponía de software para estimar los estadísticos de ajuste en el análisis factorial exploratorio y el análisis de componentes principales, pero recientes avances en el desarrollo de técnicas analíticas han permitido que tales índices se hayan incluido en programas de análisis factorial exploratorio como el programa FACTOR de Lorezno-Seva y Ferrando (2018).

Más recientemente, el desarrollo de estadísticos para testar la equivalencia (Yuan, et al., 2015; Jiang, Mai y Yuan, 2017) ha permitido añadir a los anteriores estadísticos de bondad de ajuste nuevos estadísticos que informan a los investigadores sobre el tamaño de una potencial especificación inadecuada (*misspecification*) y, en consecuencia,



permitirles establecer el grado de equivalencia del modelo a través de diferentes grupos y/o condiciones.

Hasta donde llega nuestro conocimiento, este es el primer estudio donde se pone a prueba la invarianza de un test de EF, que produce puntuaciones quasi-ipsativas, con independencia de la medida, y que ha sido examinada en condiciones de respuesta honesta y bajo instrucciones de faking. Además, se han empleado dos tipos de diseño experimental, entre grupos e intra-grupo, porque se ha demostrado que la naturaleza del diseño tiene efectos poderosos sobre la magnitud de las diferencias en las puntuaciones en condiciones de respuesta honesta y bajo instrucciones de faking. Por tanto, esa sería una primera contribución de este estudio.

La invarianza o equivalencia de la medida establece que las medidas en dos o más grupos diferentes o en dos o más condiciones distintas son comparables (Millsap, 2011; Jiang et al., 2017). En el presente caso, el examen de la equivalencia de la media se hizo con el propósito de determinar si la medida de los Big Five cambia en función de que la respuesta al cuestionario se haga con instrucciones de respuesta honesta o instrucciones de faking. Si la evidencia muestra que no existe equivalencia o invarianza de la medida, tal evidencia indicaría que las condiciones de respuesta pueden modificar o alterar la medida de los constructos evaluados y, por ello, no se estaría midiendo el mismo atributo o constructo en tales condiciones. Esta cuestión es particularmente importante en el caso de las medidas de personalidad empleadas en procesos de selección de personal, ya que existe la posibilidad de que los sujetos voluntariamente distorsionen sus respuestas, presentando una imagen más positiva de sí mismos. Al propio tiempo, el desarrollo de los inventarios de personalidad se hace de modo típico con muestras que no están motivadas para falsear o distorsionar sus respuestas.

Este es el primer estudio que ha analizado el efecto del faking sobre la invarianza o equivalencia de constructo de las medidas de personalidad de EF. En ese sentido las hipótesis que se pusieron a prueba señalaban que en las medidas de elección EF en condición faking no afectaba a la estructura factorial. Asimismo, se planteó un análisis más específico de los efectos del faking sobre la estructura factorial en una muestra intra-grupo, dado que se ha mostrado que es diseño más afectado por el faking.

Los resultados de este estudio apoyan la robustez de la estabilidad de la estructura de cinco factores en los cuestionarios de personalidad de EF ante el faking. Los resultados del análisis factorial exploratorio han mostrado en todos los casos una estructura factorial de cinco componentes. Si bien el orden de factores fue diferente en cada una de las

matrices de cargas rotadas, los índices de ajuste obtenidos en estos análisis mostraron el buen ajuste del modelo a los datos.

Asimismo, los resultados del análisis confirmatorio han mostrado, tanto en la condición honesta como en la condición faking de las dos muestras, índices de ajuste aceptables que permiten afirmar que los datos se ajustan al modelo de cinco factores propuesto y, por tanto, que la estructura factorial se mantiene estable.

Además, se han analizado los coeficientes de congruencia factorial con el objetivo de obtener un indicador más robusto de las similitudes entre las estructuras factoriales. Los resultados no dejan lugar a dudas, los valores obtenidos muestran la robustez de la estructura factorial en los cuestionarios de EF tanto en condiciones honestas como en condiciones faking incluso en diseños experimentales intra-grupo. Es decir, cuando se usan cuestionarios de EF, la estructura factorial no se ve afectada por el faking.

#### IMPLICACIONES PRÁCTICAS

Los resultados encontrados tienen una importante implicación para la práctica profesional y es que se ha demostrado que los cuestionarios de personalidad de EF son robustos contra los efectos del faking sobre la validez de constructo por lo tanto se tratan de una herramienta válida para evaluar la personalidad. En consecuencia, se recomienda su uso en los procesos de selección y evaluación de personal, ya que los cuestionarios de EF han demostrado ser una herramienta útil para reducir los efectos del faking sobre la estructura factorial.

#### LIMITACIONES DEL ESTUDIO

También cabe señalar que este estudio presenta también ciertas limitaciones que se deben tener en cuenta. La primera de ellas es que se trata de un estudio llevado a cabo en un contexto experimental de laboratorio con una muestra muy específica, estudiantes universitarios. por lo que los datos pueden estar afectados por las características particulares de la muestra. En ese sentido sería conveniente ampliar esta investigación a un contexto real (candidatos vs empleados) para obtener un mayor apoyo y evidencia a los resultados obtenidos. Otro punto a tener en cuenta es que la muestra intra-grupo formaba parte de la muestra total analizada. Es decir, los análisis no se han llevado a cabo con dos muestras totalmente independientes (inter-grupos vs intra-grupo) debido al

reducido tamaño de la muestra del diseño inter-grupos en la condición faking. No obstante, dada la concordancia en los resultados de los análisis de las muestras analizadas, los resultados podrían ser muy similares a los obtenidos.

Finalmente, este estudio ha llevado a cabo con una medida de EF quasi-ipsativa exclusivamente, por lo que resultaría de interés examinar si estos resultados se replicarían con los otros dos formatos de cuestionario de EF: ipsativos y normativos.





## ESTUDIO 3

EFECTOS DEL FAKING SOBRE LA VALIDEZ  
PREDICTIVA DE LOS CUESTIONARIOS DE  
PERSONALIDAD DE ELECCIÓN FORZOSA QUASI-  
IPSATIVOS



## INTRODUCCIÓN

El objetivo del tercer estudio de esta tesis doctoral es examinar si el faking afecta a la validez predictiva de los cuestionarios de EF quasi-ipsativos. Como hemos descrito en el capítulo 5, los meta-análisis de Salgado (2017), Salgado y Táuriz (2014) y Salgado et al. (2015) han mostrado que las medidas de personalidad de EF quasi-ipsativos son un instrumento más válido para predecir el desempeño ocupacional y el desempeño académico que las medidas de personalidad SS. Los resultados de los meta-análisis de Salgado y colaboradores (Salgado y Táuriz, 2014; Salgado, et al., 2015) indicaron que los cuestionarios de EF quasi-ipsativos eran mejores predictores del desempeño que los cuestionarios EF ipsativos o normativos. Más específicamente, dichos meta-análisis encontraron que el factor conciencia evaluado con cuestionarios de EF quasi-ipsativos es el mejor predictor del desempeño ocupacional y académico, dentro del modelo de los Cinco Grandes, con un tamaño de validez semejante al de los mejores instrumentos de predicción.

Los resultados de Salgado (2017, Salgado y Táuriz, 2014; Salgado, et al., 2015) han sido replicados recientemente por Fisher, Robie, Christiansen, Speer y Schneider (2019) en un meta-análisis a pequeña escala y por el estudio de Lee et al. (2018), quienes también encontraron que las medidas de personalidad de EF quasi-ipsativas son el instrumento de evaluación de la personalidad más válido para predecir el desempeño.

No obstante, ninguno de los estudios publicados hasta el momento ha examinado la validez predictiva de las medidas de personalidad de EF quasi-ipsativas en condiciones experimentales de inducción de faking. Por ello, es una cuestión pendiente de investigación si el faking inducido experimentalmente puede afectar y de qué modo a la validez predictiva de las medidas de personalidad de EF quasi-ipsativas.

De acuerdo con la teoría psicométrica de los efectos del faking (Salgado, 2016), las consecuencias del mismo serían dos. Por un lado, un aumento de la puntuación media y, por otro, una reducción de la varianza (desviación típica). Dichos efectos conducirían a su vez a una disminución de la fiabilidad y a una disminución de la validez predictiva. Por otra parte, los resultados del meta-análisis y los análisis factoriales descritos en los estudios 1 y 2 de esta tesis han puesto de manifiesto la robustez de las medidas de personalidad de EF quasi-ipsativas ante los efectos del faking. Por un lado, se pudo comprobar que los efectos del faking eran considerablemente menos importantes sobre la puntuación media de los grupos de evaluados (esto es, existía resistencia al faking). Por



otro, se pudo determinar que los cuestionarios de EF quasi-ipsativos mostraban un elevado grado de invarianza de medida (es decir, equivalencia de medida), tanto en condiciones de respuesta honesta como bajo instrucciones de inducción del faking.

Considerando los tres tipos de hallazgos de la investigación: (a) una validez predictiva más elevada de los cuestionarios de EF quasi-ipsativos; (b) una mayor resistencia al faking de esta modalidad de cuestionarios y; (c) una adecuada invarianza de la medida en condiciones honestas y de faking, puede esperarse que los efectos negativos del faking sobre la validez predictiva de los cuestionarios de EF quasi-ipsativos sean menores o no existan.

El tercer estudio de esta tesis está, por tanto, dirigido a comprobar esta posibilidad, poniendo a prueba la capacidad predictiva de un cuestionario de EF quasi-ipsativo para predecir el desempeño académico (evaluado a través de un cuestionario autoinforme) y los resultados académicos (medido a través de las calificaciones académicas oficiales) de estudiantes de universidad.

Teniendo en cuenta que el factor conciencia es el mejor predictor de los resultados académicos, como han encontrado los meta-análisis citados anteriormente, se plantea la siguiente hipótesis:

*H1: El factor de conciencia medido con un cuestionario de EF quasi-ipsativo predice el desempeño académico y los resultados académicos tanto en condiciones de respuesta honesta como de respuesta bajo instrucciones de faking.*

Por otra parte, la investigación previa, incluyendo los meta-análisis, no han examinado si los factores de personalidad, particularmente el factor de conciencia, predicen de igual modo el desempeño académico cuando éste es evaluado mediante escalas de valoración o mediante los resultados académicos, es decir, mediante las calificaciones. Los resultados producidos en otros ámbitos de investigación, por ejemplo, en el de las capacidades cognitivas (Salgado y Moscoso, 2019), han mostrado que los coeficientes de validez predictiva son diferentes cuando el desempeño se evalúa mediante escalas de valoración y mediante datos de resultados (p. ej., producción, ventas). En el caso de las capacidades cognitivas, la validez es más alta en las escalas de valoración. Además, investigaciones meta-analíticas han encontrado que la relación entre las medidas del desempeño basadas en datos y las medidas del desempeño basadas en juicios es de moderada magnitud (alrededor de .40). A este respecto, McDaniel, Whetzel, Schmidt y Maurer (1994) sugirieron que distinguir entre el tipo de medidas del criterio es relevante

en los estudios de validez predictiva, porque el tamaño de la validez habitualmente varía según el tipo de criterio y, por ello, se deberían realizar análisis separados para identificar potenciales diferencias entre los criterios.

En consecuencia, teniendo en cuenta los resultados que indican que el tipo de medida del desempeño puede ser una variable moderadora de la validez predictiva, se plantean las siguientes hipótesis:

*H2: La validez predictiva del factor de personalidad conciencia evaluado con cuestionarios de EF quasi-ipsativos es mayor cuando el desempeño se evalúa mediante escalas de valoración que cuando se emplean medidas de los resultados (p.ej., promedio de las calificaciones).*

*H3: Los efectos del faking sobre la validez predictiva del factor de conciencia evaluado con cuestionarios de EF quasi-ipsativos son independientes del modo en que se evalúe el desempeño (mediante escalas de valoración o con medidas de los resultados).*

Por tanto, el examen de estas dos hipótesis será otra contribución única a la literatura de esta tesis.

Un último objetivo es poner a prueba dos de las predicciones derivadas de la teoría psicométrica del faking (Salgado, 2016). Esta teoría sostiene que el faking produce una disminución de la fiabilidad y de la validez predictiva de las medidas de personalidad, a consecuencia de un incremento en la puntuación media y una reducción de la variabilidad (menor desviación estándar). Por ello, las dos últimas hipótesis de este estudio son:

*H4: El faking produce una reducción de la fiabilidad de las medidas de personalidad que puede estimarse comparando la fiabilidad en condiciones de respuesta honesta y de respuesta bajo instrucciones de faking.*

*H5: El faking produce restricción en el rango en las medidas de personalidad que puede observarse cuando se compara la variabilidad en condiciones de respuesta honesta y de respuesta en condiciones de faking.*

Hasta el momento presente ningún estudio ha investigado las predicciones de la teoría psicométrica del faking en cuestionarios de personalidad de EF quasi-ipsativos. Por ello, el examen de este objetivo será también una contribución única de este estudio.

## MÉTODO

### 1. MUESTRA

La muestra de este estudio es una parte de la muestra empleada en el estudio 2 de esta tesis. En total, 939 estudiantes procedentes de diversas titulaciones de la Universidad de Santiago de Compostela participaron en esta investigación. El 30.04 % de los sujetos fueron hombres mientras que el 69.96% eran mujeres. La edad promedio de la muestra fue de 21.62 ( $SD = 3.90$ ), el 22.58% tenía entre 17 y 20 años de edad, el 69.12% se situaba entre los 20 y los 25 años y tan sólo un 8.30% era mayores de 25 años.

Respecto a las carreras de procedencia, la mayoría de los participantes eran estudiantes del área de ciencias sociales (48.88%) seguido por ciencias de la salud (32.27%) y ciencias puras (11.71%); el 5.32% estudiaba una carrera de área de humanidades y menos del 2% procedían de ingenierías. Asimismo, la mayor parte de la muestra (97.34%) estudiaba o había finalizado recientemente los estudios de primer ciclo y sólo un 1.92 % estaban cursando un máster y menos de 1% estudios de doctorado en el momento de participar en el estudio. En la Tabla 47 se resumen las características de la misma.

Tabla 47. Características Demográficas de la Muestra ( $N = 939$ )

	<i>N</i>	%
Sexo		
<i>Mujeres</i>	657	69.96
<i>Hombres</i>	282	30.04
Edad		
<i>17-20</i>	212	22.58
<i>20-25</i>	649	69.12
<i>≥25</i>	78	8.30
Nivel de estudios		
<i>Primer ciclo</i>	914	97.34
<i>Segundo ciclo</i>	18	1.92
<i>Tercer ciclo</i>	7	0.74
Área de estudio		
<i>Ciencias sociales</i>	459	48.89
<i>Ciencias</i>	110	11.71
<i>Ciencias de la salud</i>	303	32.27
<i>Humanidades</i>	50	5.32
<i>Ingenierías</i>	17	1.81

*Nota.* *N* = número de sujetos; Primer ciclo = estudios de diplomatura, licenciatura o grado; Segundo ciclo = estudios de máster; Tercer ciclo = estudios de doctorado.

Los métodos utilizados para el reclutamiento fueron los descritos en el estudio 2: (a) el reclutamiento en las aulas y (b) la solitud de participantes en un estudio experimental a través de la publicación de anuncios informativos en las diferentes facultades y edificios administrativos de la Universidad de Santiago de Compostela. Asimismo, en este caso, se incentivó la participación informando que cada participante recibiría una compensación económica (entre 10 y 15 euros) por participar voluntariamente en el estudio. Finalmente, el 19.49% de los participantes fueron reclutados en las aulas y el 80.51% participaron voluntariamente en el estudio experimental.

Para poder participar en el estudio se exigió que los participantes cumpliesen dos condiciones: (a) ser estudiante universitario y (b) entregar una copia del expediente académico en la que apareciese la nota de las materias superadas y la nota media del expediente. Asimismo, antes de cubrir las pruebas, los participantes fueron informados de que la información recogida sería tratada de modo totalmente confidencial y que se usaría únicamente con fines de investigación. Además, se garantizó el anonimato en el tratamiento de los datos. Todos los participantes firmaron un documento de consentimiento informado.

El 52.18% ( $N = 490$ ) de la muestra participó en un diseño intra-grupo en el que los participantes contestaron al cuestionario de modo honesto y bajo condiciones de faking, mientras que el restante 47.82% ( $N = 449$ ) contestaron al cuestionario de personalidad sólo en la condición honesta.

## 2. INSTRUMENTOS DE MEDIDA

### *Personalidad*

*QI5F\_tri*. Para evaluar la personalidad se utilizó el mismo cuestionario de personalidad que se empleó en el estudio 2 de esta tesis, el cuestionario de personalidad de EF quasi-ipsativo QI5F-tri (Salgado, 2014). Nos remitimos a la descripción de este cuestionario que presentamos en dicho estudio.

### *Resultados Académicos*

*CAP*. La evaluación de los resultados académicos se ha realizado a través de la calificación académica promedio (CAP) del expediente académico (en inglés *grade point average*, *GPA*) de los participantes, o lo que es lo mismo, la puntuación media del

conjunto de materias superadas en el momento de participar en el estudio. En el sistema universitario español esta puntuación es el resultado de dividir el sumatorio resultante de multiplicar los créditos de cada materia por la calificación obtenida en dicha materia entre el número total de créditos aprobados.

#### *Desempeño de tarea*

*CDTE.* La segunda medida utilizada para evaluar el desempeño académico de tarea fue la escala CDTE desarrollada por Salgado (2010). Se trata de una medida de autoinforme compuesta por 30 ítems que evalúan el logro de tareas, la implicación y el cumplimiento de deberes como estudiantes. El sujeto debe indicar, a través de una escala tipo Likert de cinco puntos, con qué frecuencia realiza cada comportamiento descrito. La fiabilidad por consistencia interna de esta medida fue de  $\alpha = .89$ .

#### *Desempeño Contextual*

*CDCE.* En este estudio se utilizó la escala CDCE (Salgado, 2010) para evaluar el desempeño contextual (también denominado desempeño cívico), como otro importante criterio de éxito académico. Esta escala tiene 30 ítems que evalúan las siguientes conductas: apoyo personal (12 ítems), apoyo organizacional (9 ítems) e iniciativa consciente (9 ítems). Los participantes tuvieron que indicar la frecuencia con la que realizan las conductas descritas en cada pregunta utilizando una escala tipo Likert de 1 (*nunca*) a 5 (*siempre*). puntos. La fiabilidad encontrada para esta medida fue de  $\alpha = .86$ .

#### *Conductas Contraproductivas Académicas*

*CDAN.* Respecto a la medida de conductas contraproductivas académicas se utilizó la escala CDAN de Salgado (2010). Esta escala se compone de 30 ítems que evalúan las siguientes conductas: copiar, uso inapropiado de recursos, absentismo, incumplimiento de normas y bajo esfuerzo. Los sujetos tienen que marcar su grado de acuerdo con cada ítem en una escala del 1 al 5, en donde 1 significa *nunca* y 5 significa *siempre*. El alfa de Cronbach obtenida para este factor fue de .89.

### 3. PROCEDIMIENTO

En este estudio, la duración de cada sesión evaluativa fue de aproximadamente dos horas. Para el caso de la medida de personalidad, se siguió exactamente el mismo

procedimiento descrito en el estudio 2 para responder a este cuestionario bajo las dos condiciones experimentales evaluadas: honesta y faking. Esto es, en la condición honesta los sujetos siguieron las instrucciones facilitadas por el propio cuestionario, mientras que en la condición faking se realizaron ligeras modificaciones en las instrucciones del instrumento para incitar a los participantes a cometer faking. Concretamente, se añadió el siguiente párrafo: *“Al responder imagínese que está en la última parte del proceso de selección para un puesto muy atractivo, dado que le ofrece grandes posibilidades de desarrollar su carrera profesional. Usted quiere conseguir ese puesto. Para ello, debe contestar a este test tratando de dar la mejor imagen posible para conseguir ese puesto.”*

En el diseño intra-grupo se presentaron de modo contrabalanceado las instrucciones para las dos condiciones experimentales (honestas y faking), con el objetivo de controlar los posibles sesgos debidos al orden de presentación de las condiciones. De tal modo que el 47.55% ( $N=233$ ) de los participantes recibieron primero las instrucciones para responder honestamente y posteriormente contestaron al test en condiciones de faking, mientras que el 52.45% ( $N=257$ ) respondieron primero al cuestionario con instrucciones que inducían al faking y después contestaron honestamente. Los participantes tardaron 45 minutos aproximadamente en responder al cuestionario de personalidad en cada condición.

Respecto al procedimiento para las medidas de desempeño, en las tres pruebas (CDTE, CDCE, Y CDAN) se les indicó a los participantes que debían seguir las instrucciones propias del cuestionario:

*“Por favor, indique hasta qué punto ha llevado a cabo las conductas y acciones descritas más abajo, que pueden haberle ocurrido en su entorno académico. Use la escala que aparecen seguidamente, empleando los números 1 a 5: 1 = nunca; 2 = casi nunca; 3 = a veces; 4 = casi siempre; 5 = siempre.”*

Asimismo, se les pidió que fuesen totalmente honestos en sus respuestas, reiterando que se trataría la información de modo anónimo y sin ninguna repercusión para ellos a nivel académico. Los sujetos tardaron entre 5 y 10 minutos en responder a cada una de las escalas de desempeño. Los participantes en el diseño intra-grupo respondieron a estas escalas entre las dos condiciones experimentales (honestas y faking) de la medida de personalidad para separar en el tiempo las respuestas de los sujetos bajo unas y otras instrucciones. Con esto se pretendía conseguir que los sujetos no pudieran apoyarse en lo contestado en la primera condición del test de personalidad para responder en la segunda.



## RESULTADOS

A continuación, se presentan los resultados obtenidos en el análisis de correlaciones entre la personalidad y el conjunto de criterios de desempeño examinados en las condiciones honesta y faking. Cada una de las tablas que se presentan incluye, en primer lugar, las medias y desviaciones típicas obtenidas para cada una de las variables y, a continuación, las correlaciones observadas entre las variables de personalidad, las medidas de desempeño de tarea, contextual y de deshonestidad académica, respectivamente. Primero se presentan los resultados obtenidos para la condición honesta, en la que se han examinado los resultados a través de tres muestras: (a) la muestra total de los sujetos participantes en el estudio; (b) la muestra que participó en el diseño inter-grupos y (c) la muestra intra-grupo. Tras estos resultados se presentan los datos obtenidos en la condición faking para la muestra intra-grupo. Siguiendo a cada tabla de resultados, aparece la tabla en la que se indican los tamaños muestrales empleados para calcular cada una de las relaciones descritas.

En la Tabla 48A se presentan las correlaciones observadas obtenidas entre las variables del estudio analizadas para la muestra total en la condición honesta. Como puede observarse, los resultados muestran correlaciones significativas entre la mayoría de los factores de personalidad y los criterios de desempeño, con la excepción del factor apertura a la experiencia que no correlacionó significativamente con ningún criterio analizado.

El factor estabilidad emocional se relaciona de forma negativa y significativa con el desempeño de tarea (CDTE;  $r = -.15, p < .01$ ) y contextual (CDCE;  $r = -.13, p < .01$ ), obteniéndose tamaños del efecto muy semejantes. Esto indica que las personas más estables emocionalmente tienden a tener peor desempeño académico y contextual. Este factor también correlaciona de modo negativo y significativo con las conductas contraproductivas académicas, de tal modo que cuanto más estable emocionalmente sea la persona menos propensa será también tener este tipo de comportamientos, aunque en este caso el tamaño del efecto es más pequeño ( $r = -.09, p < .01$ ). No se ha encontrado una relación significativa entre este factor y la CAP ( $r = -.06$ ).

Con respecto al factor extraversión, los resultados muestran que predice el desempeño de tarea. La correlación obtenida con la medida CDTE es negativa con un valor de  $r = -.10$  ( $p < .01$ ), lo que indica que a mayor puntuación en extraversión peor desempeño de tarea. Además, también se relaciona de modo positivo y significativo con



las conductas académicas contraproductivas con una correlación de  $r = .15$  ( $p < .01$ ), es decir, cuanto más extrovertida sea la persona más propensa será a cometer este tipo de conductas. Las correlaciones obtenidas entre este factor y los resultados académicos medidos a través de la CAP y el desempeño contextual han sido muy bajas y no significativas, con valores de  $r = -.04$  y  $r = .03$ , respetivamente.

Apertura a la experiencia, ha sido el único factor que no ha correlacionado de modo significativo con ninguno de los criterios de desempeño analizados, siendo en todos los casos tamaños del efecto inferiores a .07. Obteniéndose correlaciones negativas con los criterios de desempeño positivos (CAP,  $r = -.02$ ; CDTE,  $r = -.04$ ; CDCE,  $r = -.03$ ) y una correlación positiva ( $r = .07$ ) con el criterio de desempeño contraproductivo (CDAN) evaluado.

El factor amigabilidad correlacionó de forma negativa y significativa exclusivamente con la calificación académica promedio ( $r = -.14$ ,  $p < .01$ ), lo que viene a indicar que aquellos sujetos que obtengan puntuaciones más altas en amigabilidad tendrán un peor desempeño en el ámbito académico. También se ha encontrado una correlación baja pero significativa entre este factor y el desempeño contextual de  $r = .07$  ( $p < .05$ ), lo que indica que amigabilidad predice este tipo de conductas. Los resultados con las restantes variables criterio examinadas muestran correlaciones muy bajas y no significativas.

Tabla 48 A. Medias, Desviaciones Típicas y Correlaciones Observadas entre las Variables de la Muestra Total en la Condición Honesta

	<i>M</i>	<i>SD</i>	EE	EX	AE	A	C	CAP	CDTE	CDCE	CDAN
1. EE	22.66	6.60	-								
2. EX	27.98	8.03	-.03	-							
3. AE	28.49	9.57	-.11**	-.19**	-						
4. A	29.58	6.80	-.02	-.20**	-.15**	-					
5. C	26.36	8.23	-.19**	-.44	-.17**	-.13**	-				
6. CAP ( <i>z</i> )	6.98 (.04)	.88 (.99)	-.06	-.04	-.02	-.14**	.24**	-			
7. CDTE	102.06	15.01	-.15**	-.10**	-.04	-.05	.38**	.27**	-		
8. CDCE	108.64	12.63	-.13**	.03	-.03	.07*	.17**	.16**	.57**	-	
9. CDAN	61.41	13.31	-.09**	.15**	.07	-.03	-.31**	-.25**	-.56**	-.37**	-

*Nota.* *M* = puntuación promedio de la variable; *SD* = desviación típica de las puntuaciones de la variable; EE = estabilidad emocional; EX = extraversión; AE = apertura a la experiencia; A = amigabilidad; C = conciencia; CAP = calificación académica promedio; CDTE = desempeño de tarea; CDCE = desempeño contextual; CDAN = desempeño académico contraproducente \* $p < .05$ . \*\* $p < .01$ .

Tabla 48B. Tamaños Muestrales Utilizados para el Cálculo de las Correlaciones Observadas entre las Variables de la Muestra Total en la Condición Honesta

	<i>M</i>	<i>SD</i>	EE	EX	AE	A	C	CAP	CDTE	CDCE	CDAN
1. EE	22.66	6.60	-								
2. EX	27.98	8.03	939	-							
3. AE	28.49	9.57	939	939	-						
4. A	29.58	6.80	939	939	939	-					
5. C	26.36	8.23	939	939	939	939	-				
6. CAP ( <i>z</i> )	6.98 (.04)	.88 (.99)	777	777	777	777	777	-			
7. CDTE	102.06	15.01	803	803	803	803	803	769	-		
8. CDCE	108.64	12.63	794	794	794	794	794	765	859	-	
9. CDAN	61.41	13.31	799	799	799	799	799	766	825	816	-

*Nota.* *M* = puntuación promedio de la variable; *SD* = desviación típica de las puntuaciones de la variable; EE = estabilidad emocional; EX = extraversión; AE = apertura a la experiencia; A = amigabilidad; C = conciencia; CAP = calificación académica promedio; CDTE = desempeño de tarea; CDCE = desempeño contextual; CDAN = desempeño académico contraproducente.

Por lo que respecta al factor conciencia demuestra ser un robusto predictor del desempeño. En todos los casos se han encontrado correlaciones significativas entre este factor y los criterios evaluados. Así, conciencia correlaciona de modo positivo y significativo tanto con la medida de los resultados (calificación académica promedio) como con las tres medidas de desempeño académico analizadas. En relación con la calificación académica promedio se obtuvo una correlación de  $r = .24$  ( $p < .01$ ). Con respecto al desempeño se encontró una correlación de  $r = .38$  ( $p < .01$ ) con el desempeño de tarea, de  $r = .17$  con el desempeño contextual y de  $r = -.31$  con la medida de conductas contraproductivas. Estos resultados apoyan la Hipótesis 1 en lo que se refiere a las respuestas en condiciones honestas y están de acuerdo con los resultados del meta-análisis de Salgado y Táuriz (2014).

Como hemos descrito en el procedimiento estas relaciones se han analizado con el conjunto de datos obtenidos mediante dos diseños experimentales diferentes: un diseño inter-grupos y un diseño intra-grupo. Con objetivo de tener una visión más precisa de la relación entre las variables en cada tipo de diseño, a continuación, se desglosan los resultados en función de este criterio para realizar una comparativa entre los resultados de la muestra total y cada uno de los diseños.

En la Tabla 49A y 50A se presentan las correlaciones obtenidas en la condición honesta para la muestra independiente (inter-grupos) y para el diseño intra-grupo, respectivamente. Los datos se presentan en el mismo orden que para la muestra total.

En términos generales, los resultados obtenidos son muy semejantes entre estos diseños y los obtenidos para la muestra total en la condición honesta. Los resultados muestran que independientemente del diseño experimental la personalidad evaluada con medidas de EF quasi-ipsativas predice los resultados académicos y los criterios de desempeño, aunque hay diferencias en la magnitud de las correlaciones.

Con respecto a los resultados del diseño inter-grupos (Tabla 49A), puede observarse que las correlaciones significativas obtenidas son ligeramente superiores a las encontradas en la muestra total. El factor estabilidad emocional correlaciona significativamente con los mismos criterios que la muestra total, desempeño de tarea (CDTE) y desempeño contextual, pero en este caso con un valor  $r = -.18$  ( $p < .01$ ), superior al obtenido en la muestra total para ambos casos. La excepción es la correlación obtenida con la escala CDAN, a pesar de que el valor de la correlación es prácticamente el mismo que para la muestra total ( $r = .10$ ), en este diseño esta relación no es significativa.

El factor extraversión muestra correlaciones significativamente más altas en esta muestra con los resultados académicos (CAP), con la medida de desempeño de tarea CDTE, ( $r = -.16, p < .01$ ) y con la medida de conductas académicas contraproductivas CDAN ( $r = .25, p < .01$ ), que en la muestra total.

Los factores apertura a la experiencia y amigabilidad no han obtenido correlaciones significativas con ninguno de los criterios de desempeño evaluados. En el caso de factor apertura a la experiencia estos resultados siguen la línea de los obtenidos para la muestra total. De hecho, no se han producido cambios sustanciales en los valores de las correlaciones obtenidas. No obstante, el factor amigabilidad para esta condición no predice ninguno de los criterios de desempeño de tarea y contextual con los que correlacionaba significativamente en la muestra total.

Por último, conciencia, en la condición inter-grupos, muestra correlaciones significativas con todas las variables criterio utilizadas en este estudio, lo que está de acuerdo con lo planteado en la Hipótesis 1 y también con los resultados de los meta-análisis previos. En casi todos los casos los valores de tamaño del efecto *son* más altos que los obtenidos en la muestra total. Muestra correlaciones de  $r = .34$  ( $p < .01$ ) y  $r = .38$  ( $p < .01$ ) con las variables CAP y CDTE, una  $r = .15$  ( $p < .01$ ) como predictor del desempeño contextual y, por último, una correlación de  $r = -.33$  ( $p < .01$ ) con el criterio de conductas académicas contraproductivas. Estos resultados muestran de nuevo que conciencia es el factor que mejor predice el desempeño y los resultados académicos.

Tabla 49 A. Medias, Desviaciones Típicas y Correlaciones Observadas entre las Variables del Diseño Inter-Grupos en la Condición Honesta

	<i>M</i>	<i>SD</i>	EE	EX	AE	A	C	CAP	CDTE	CDCE	CDAN
1. EE	22.83	6.70	-								
2. EX	28.69	8.05	-.03	-							
3. AE	26.87	9.61	-.09*	-.12*	-						
4. A	29.71	6.96	-.03	-.23**	-.16**	-					
5. C	26.38	7.98	-.19**	-.46**	-.16**	-.09	-				
6. CAP ( <i>z</i> )	6.95 (.04)	.85 (.97)	-.07	<b>-.16**</b>	-.02	-.10	<b>.34**</b>	-			
7. CDTE	100.23	15.20	<b>-.18**</b>	<b>-.16**</b>	.03	-.01	<b>.38**</b>	.21**	-		
8. CDCE	109.50	12.24	<b>-.18**</b>	.00	-.01	.10	<b>.15**</b>	.14*	.59**	-	
9. CDAN	62.18	12.79	.10	<b>.25**</b>	.01	-.10	<b>-.33**</b>	-.22**	-.57**	-.34**	-

*Nota.* *M* = puntuación promedio de la variable; *SD* = desviación típica de las puntuaciones de la variable; EE = estabilidad emocional; EX = extraversión; AE = apertura a la experiencia; A = amigabilidad; C = conciencia; CAP = calificación académica promedio; CDTE = desempeño de tarea; CDCE = desempeño contextual; CDAN = desempeño académico contraproducente. \* $p < .05$ . \*\* $p < .01$ .

Tabla 49B. Tamaños Muestrales Utilizados para el Cálculo de las Correlaciones Observadas entre las Variables del Diseño Inter-grupos en la Condición Honesta

	<i>M</i>	<i>SD</i>	EE	EX	AE	A	C	CAP	CDTE	CDCE	CDAN
1. EE	22.83	6.70	-								
2. EX	28.69	8.05	449	-							
3. AE	26.87	9.61	449	449	-						
4. A	29.71	6.96	449	449	449	-					
5. C	26.38	7.98	449	449	449	449	-				
6. CAP ( <i>z</i> )	6.95 (.04)	.85 (.97)	338	338	338	338	338	-			
7. CDTE	100.23	15.20	331	331	331	331	331	309	-		
8. CDCE	109.50	12.24	326	326	326	326	326	308	326	-	
9. CDAN	62.18	12.79	330	330	330	330	330	309	325	330	-

*Nota.* *M* = puntuación promedio de la variable; *SD* = desviación típica de las puntuaciones de la variable; EE = estabilidad emocional; EX = extraversión; AE = apertura a la experiencia; A = amigabilidad; C = conciencia; CAP = calificación académica promedio; CDTE = desempeño de tarea; CDCE = desempeño contextual; CDAN = desempeño académico contraproducente.



Por su parte, los resultados obtenidos en la condición honesta para el diseño intra-grupo (Tabla 50A) muestran, en general, correlaciones ligeramente más bajas que las obtenidas para el diseño inter-grupos y la muestra total. En este diseño, el factor estabilidad emocional predice de modo significativo, exclusivamente, el desempeño académico (CDTE) con una correlación negativa de  $r = -.13$  ( $p < .01$ ), siendo muy bajas las correlaciones obtenidas con las restantes medidas de desempeño.

En el caso del factor extraversión, en este diseño experimental no se obtienen correlaciones significativas entre este factor y los criterios evaluados. Por lo tanto, para este caso, y contrariamente a los resultados obtenidos en las otras muestras, extraversión no resulta ser predictor de ningún criterio de desempeño.

Lo contrario, ocurre con el factor apertura a la experiencia, que en este diseño predice el desempeño académico,  $r = -.10$  ( $p < .01$ ), y las conductas desviadas en el contexto académico,  $r = .12$  ( $p < .01$ ).

El factor amigabilidad, correlaciona de modo significativo y negativo con el criterio de desempeño académico GPA, obteniendo una correlación de  $r = -.17$  ( $p < .01$ ). Dato ligeramente superior al obtenido en la muestra total. Sin embargo, no se obtiene correlaciones significativas con las restantes medidas de desempeño, siendo en todos los casos valores muy próximos a cero.

El factor conciencia, nuevamente, es el único que obtiene correlaciones significativas con todas las variables criterio analizadas. Las correlaciones son positivas y significativas con CAP ( $r = .17$ ,  $p < .01$ ), el desempeño de tarea CDTE ( $r = .38$ ,  $p < .01$ ), el desempeño contextual CDCE ( $r = .18$ ,  $p < .01$ ), y muestra una relación negativa y significativa como predictor de las conductas contraproductivas académicas CDAN, con una  $r = -.30$  ( $p < .01$ ). Estos datos muestran la robustez del factor conciencia como predictor del desempeño cuando se utilizan medidas de EF quasi-ipsativas y, nuevamente, apoyan parcialmente la Hipótesis 1.

En resumen, los resultados obtenidos en la condición honesta para la muestra total y los dos diseños experimentales muestran que la personalidad, en general, y el factor conciencia, en particular, son importantes predictores de múltiples criterios de desempeño cuando se usan cuestionario de EF quasi-ipsativos.

Tabla 50 A. Medias, Desviaciones Típicas y Correlaciones Observadas entre las Variables del Diseño Intra-Grupo en la Condición Honesta

	<i>M</i>	<i>SD</i>	EE	EX	AE	A	C	CAP	CDTE	CDCE	CDAN
1. EE	22.49	6.52	-								
2. EX	27.34	7.96	-.03	-							
3. AE	29.97	9.30	-.13**	-.11*	-						
4. A	29.46	6.64	-.01	-.17**	-.15**	-					
5. C	26.34	8.47	-.19**	-.43**	-.18**	.16**	-				
6. CAP ( <i>z</i> )	7.01(.33)	.92(1.03)	-.06	.05	-.01	-.17**	.17**	-			
7. CDTE	102.76	14.80	-.13**	-.04	-.10*	-.08	.38**	.30**	-		
8. CDCE	107.99	12.68	-.09	.04	.04	-.05	.18**	.18**	.57**	-	
9. CDAN	61.30	13.68	.08	.08	.12**	.01	-.30**	-.24**	-.55**	-.38**	-

*Nota.* *M* = puntuación promedio de la variable; *SD* = desviación típica de las puntuaciones de la variable; EE = estabilidad emocional; EX = extraversión; AE = apertura a la experiencia; A = amigabilidad; C = conciencia; CAP = calificación académica promedio; CDTE = desempeño de tarea; CDCE = desempeño contextual; CDAN = desempeño académico contraproducente. \* $p < .05$ . \*\* $p < .01$ .

Tabla 50B. Tamaños Muestrales Utilizados para el Cálculo de las Correlaciones Observadas entre las Variables del Diseño Intra-grupo en la Condición Honesta

	<i>M</i>	<i>SD</i>	EE	EX	AE	A	C	CAP	CDTE	CDCE	CDAN
1. EE	22.49	6.52	-								
2. EX	27.34	7.96	490	-							
3. AE	29.97	9.30	490	490	-						
4. A	29.46	6.64	490	490	490	-					
5. C	26.34	8.47	490	490	490	490	-				
6. CAP ( <i>z</i> )	7.01(.33)	.92(1.03)	439	439	439	439	439	-			
7. CDTE	102.76	14.80	472	472	472	472	472	433	-		
8. CDCE	107.99	12.68	468	468	468	468	468	431	468	-	
9. CDAN	61.30	13.68	469	469	469	469	469	431	469	466	-

*Nota.* *M* = puntuación promedio de la variable; *SD* = desviación típica de las puntuaciones de la variable; EE = estabilidad emocional; EX = extraversión; AE = apertura a la experiencia; A = amigabilidad; C = conciencia; CAP = calificación académica promedio; CDTE = desempeño de tarea; CDCE = desempeño contextual; CDAN = desempeño académico contraproducente. \* $p < .05$ . \*\* $p < .01$ .

A continuación, se examinó si el faking afecta a la validez predictiva de estas medidas. En Tabla 51A se incluyen los resultados obtenidos para la condición faking en el diseño intra-grupo. En conjunto, los resultados muestran que la personalidad evaluada con cuestionarios de EF quasi-ipsativos predice el desempeño en condiciones de faking. Los resultados, por tanto, apoyan la  $H_1$  planteada en este estudio. No obstante, se produce una reducción de los tamaños del efecto.

Si examinamos individualmente los resultados para cada uno de los factores de personalidad, los datos muestran que el factor estabilidad emocional sigue siendo en condiciones faking predictor del desempeño académico, en base al criterio CDTE, y del desempeño contextual CDCE, con valores de  $r = -.11$  ( $p < .01$ ) y  $r = -.15$  ( $p < .01$ ), respectivamente.

El factor extraversión no muestra correlaciones significativas con ninguno de los criterios evaluados, obteniéndose resultados semejantes a los encontrados para este factor en la muestra total. En el caso del factor apertura a la experiencia, los resultados muestran que tampoco predice ningún criterio de desempeño bajo esta condición experimental, encontrándose en todos los casos correlaciones muy bajas y próximas a cero. Los resultados son muy similares a los obtenidos para este factor en la muestra total y en el diseño inter-grupos. Con el factor amigabilidad ocurre una situación semejante, ya que no se han obtenido correlaciones significativas en ninguna de las relaciones examinadas, encontrándose en todos los casos tamaños del efecto muy próximos a cero.

El factor conciencia es el único factor que predice, en condiciones faking, la práctica totalidad de los criterios de desempeño evaluados. Muestra una correlación de  $r = .11$  ( $p < .05$ ) con los resultados académicos CAP y de  $r = .16$  ( $p < .01$ ) con la escala CDTE, lo que indica que incluso en condiciones de faking, el factor conciencia predice el desempeño académico. Asimismo, se ha obtenido una correlación de  $r = -.15$  ( $p < .01$ ) con la escala CDAN, por tanto, conciencia predice la propensión de las personas a cometer conductas contraproductivas académicas también en condiciones de faking. El único criterio con el que la correlación obtenida no es significativa es con el desempeño contextual (CDCE). Si bien es cierto, que, en condiciones honestas, las correlaciones obtenidas han sido más bajas entre conciencia y este criterio en comparación con los restantes criterios de desempeño examinados. Por lo tanto, los resultados obtenidos apoyan la Hipótesis 1 en lo referente al faking.

Tabla 51 A. Medias, Desviaciones Típicas y Correlaciones Observadas entre las Variables del Diseño Intra-grupo en la Condición Faking

	<i>M</i>	<i>SD</i>	EE	EX	AE	A	C	CAP	CDTE	CDCE	CDAN
1. EE	24.51	5.53	-								
2. EX	25.80	6.41	-.04	-							
3. AE	30.94	7.52	-.12**	-.13**	-						
4. A	27.02	5.89	-.08	-.05	-.19**	-					
5. C	32.99	8.29	-.01	-.36**	-.03	-.37**	-				
6. CAP (z)	7.01(.33)	.92(1.03)	-.02	.01	.05	-.05	.11*	-			
7. CDTE	102.76	14.80	-.11*	-.00	.03	-.04	.16**	.30**	-		
8. CDCE	107.99	12.68	-.15**	.04	.02	.05	.06	.18**	.57**	-	
9. CDAN	61.30	13.68	.02	.03	-.03	.06	-.15**	-.24**	-.55**	-.38**	-

*Nota.* *M* = puntuación promedio de la variable; *SD* = desviación típica de las puntuaciones de la variable; EE = estabilidad emocional; EX = extraversión; AE = apertura a la experiencia; A = amigabilidad; C = conciencia; CAP = calificación académica promedio; CDTE = desempeño de tarea; CDCE = desempeño contextual; CDAN = desempeño académico contraproducente. \* $p < .05$ . \*\* $p < .01$ .

Tabla 51B. Tamaños Muestrales Utilizados para el Cálculo de las Correlaciones Observadas entre las Variables del Diseño Intra-grupo en la Condición Faking

	<i>M</i>	<i>SD</i>	EE	EX	AE	A	C	CAP	CDTE	CDCE	CDAN
1. EE	24.51	5.53	-								
2. EX	25.80	6.41	490	-							
3. AE	30.94	7.52	490	490	-						
4. A	27.02	5.89	490	490	490	-					
5. C	32.99	8.29	490	490	490	490	-				
6. CAP ( <i>z</i> )	7.01(.33)	.92(1.03)	439	439	439	439	439	-			
7. CDTE	102.76	14.80	472	472	472	472	472	433	-		
8. CDCE	107.99	12.68	468	468	468	468	468	431	468	-	
9. CDAN	61.30	13.68	469	469	469	469	469	431	469	466	-

*Nota.* *M* = puntuación promedio de la variable; *SD* = desviación típica de las puntuaciones de la variable; EE = estabilidad emocional; EX = extraversión; AE = apertura a la experiencia; A = amigabilidad; C = conciencia; CAP = calificación académica promedio; CDTE = desempeño de tarea; CDCE = desempeño contextual; CDAN = desempeño académico contraproducente.

Los resultados encontrados también han dado cobertura a la Hipótesis 2, que planteaba que el factor de conciencia era mejor predictor de los criterios basados en escalas de valoración que en resultados académicos. Así, los resultados en el caso de la muestra total (Tabla 48A), la validez para la CAP ( $r = .24$ ) es menor que la encontrada para el desempeño de tarea y las conductas contraproductivas ( $r = .38$  y  $r = -.31$ , respectivamente), aunque mayor que la encontrada para el desempeño contextual ( $r = .17$ ). En el caso del diseño entre grupos, la validez para predecir la CAP ( $r = .34$ ) es menor que la validez para predecir el desempeño de tarea ( $r = .38$ ), semejante a la validez para predecir el desempeño contraproductivo ( $r = .33$ ) y mayor que la validez para predecir el desempeño contextual ( $r = .15$ ). En el caso del diseño intra-grupo, la validez para predecir la CAP ( $r = .17$ ) es menor que la validez para las tres medidas de escalas de valoración, con correlaciones de .38, .18 y -.30 para CDTE, CDCE y CDAN, respectivamente. Por último, en el caso de la condición faking en el diseño intra-grupo, la validez para predecir la CAP ( $r = .11$ ) es menor que la validez para predecir el desempeño de tarea y el contraproductivo (correlaciones de .16 y -.15, respectivamente). Por tanto, en conjunto los resultados muestran un apoyo importante a la Hipótesis 2.

Sin embargo, las comparaciones anteriores no son totalmente adecuadas, ya que, por un lado, se está comparando los resultados para la CAP, que es una medida amplia (broad) del desempeño, con los resultados para las escalas que evalúan criterios facetas o subdimensiones del desempeño, por tanto, son medidas estrechas (narrow). Por otro lado, las correlaciones observadas bajo instrucciones de faking están sometidas a una potencial atenuación debida a los efectos del faking sobre la fiabilidad de las medidas de personalidad y también a una potencial subestimación debida a la potencial restricción en el rango. Una manera más adecuada de realizar la comparación es comparar los resultados para la CAP con los resultados de un compuesto de las tres facetas del desempeño académico y al propio tiempo corregir tales correlaciones por error de medida en el predictor y el criterio y restricción en el rango del predictor.

Esta comparación se ha hecho en tres pasos. En primer lugar, se realizó un análisis de componentes principales y se obtuvieron la puntuación factorial para cada sujeto de la combinación de las tres escalas. El análisis de componentes principales mostró que un único componente, con una raíz latente de 2.001 (el siguiente tenía una raíz latente de .618), explicaba el 66.7% de la varianza y las cargas factoriales eran .89, .79 y .78 para CDTE, CDCE y CDAN, respectivamente. Por tanto, un componente global del desempeño académico explica adecuadamente la relación entre las tres facetas o



subdimensiones. El segundo paso consistió en obtener la correlación entre el factor global de desempeño académico y las cinco dimensiones de personalidad, tanto en respuesta honesta como con instrucciones de faking. El tercer paso consistió en corregir las correlaciones observadas por error de medida en el predictor y el criterio y por restricción en el rango directa en el predictor, para lo que se empleó el programa VALCOR (Salgado, 1997). La Tabla 52 muestra los coeficientes de fiabilidad y restricción en el rango empleados en los análisis y las correlaciones, tanto observadas como corregidas, aparecen en la Tabla 53.

Tabla 52. Fiabilidad y Valor de Restricción en el Rango ( $u$ ) de los Big Five en Condición Honesta y Faking y Fiabilidad de las Medidas Criterio

Variable	$\alpha$ -Honesta	$\alpha$ -Faking	$u$
EE	.645	.515	.834
EX	.747	.645	.807
AE	.824	.730	.809
AM	.669	.601	.887
CO	.782	.780	.978
CAP	.830	.830	-
CDTE	.782	.782	-
CDCE	.540	.540	-
DAG	.750	.750	-
DAT	.700	.700	-

*Nota.* EE = estabilidad emocional; EX = extroversión; AE = apertura a la experiencia; AM = amigabilidad; CO = conciencia; CAP = calificación académica promedio; CDTE = desempeño de tarea; CDCE = desempeño contextual; CDAN = desempeño académico contraproductivo; DAG = desempeño académico global; DAT = desempeño académico total;  $u$  = valor de restricción en el rango.

Tabla 53. Correlaciones entre los Big Five en Condiciones Honesta y Faking con el Desempeño Global Empleando la Muestra del Diseño Intra-Grupo

Criterio	EE <sub>H</sub>	EX <sub>H</sub>	AE <sub>H</sub>	AM <sub>H</sub>	CO <sub>H</sub>	EE <sub>F</sub>	EX <sub>F</sub>	AE <sub>F</sub>	AM <sub>F</sub>	CO <sub>F</sub>
CAP – Obs	-.06	.05	-.01	-.17	.17	.02	.01	.05	-.05	.11
<b>CAP – Corr</b>	<b>-.08</b>	<b>.06</b>	<b>-.01</b>	<b>-.23</b>	<b>.21</b>	<b>.04</b>	<b>.02</b>	<b>.08</b>	<b>-.08</b>	<b>.14</b>
DAG – Obs	-.13	-.03	-.10	-.01	.35	-.11	.00	.03	-.02	.15
<b>DAG – Corr</b>	<b>-.19</b>	<b>-.04</b>	<b>-.13</b>	<b>-.01</b>	<b>.47</b>	<b>-.21</b>	<b>.00</b>	<b>.05</b>	<b>-.03</b>	<b>.20</b>

*Nota.* EE<sub>H</sub> = estabilidad emocional, condición honesta; EX<sub>H</sub> = extroversión, condición honesta; AE<sub>H</sub> = apertura a la experiencia, condición honesta; AM<sub>H</sub> = amigabilidad, condición honesta; CO<sub>H</sub> = conciencia, condición honesta; EE<sub>F</sub> = estabilidad emocional, condición faking; EX<sub>F</sub> = extroversión, condición faking; AE<sub>F</sub> = apertura a la experiencia, condición faking; AM<sub>F</sub> = amigabilidad, condición faking; CO<sub>F</sub> = conciencia, condición faking; CAP – Obs = correlación observada de los Big Five con la calificación académica promedio; CAP – Corr = correlación corregida de los Big Five con la calificación académica promedio; DAG – Obs = correlación observada de los Big Five con el desempeño académico global; DAG – Corr = correlación corregida de los Big Five con el desempeño académico global.

Los resultados reproducidos en la Tabla 53 indican que el factor de conciencia predice mejor el desempeño académico cuando está evaluado mediante escalas de valoración (juicios) que cuando está evaluado mediante resultados (calificación promedio), tanto en la condición de respuesta honesta como en la condición de respuesta con instrucciones de faking, aunque hay mayor diferencia absoluta en el primer caso. En la condición de respuesta honesta encontramos una correlación corregida de .47 para el desempeño evaluado con escalas de valoración y una correlación corregida de .21 cuando se evalúa a través de la calificación promedio. En la condición de respuesta bajo instrucciones de faking, las correlaciones corregidas fueron .20 para el desempeño medido por las escalas de valoración y .14 para la medida de resultados. Resultados semejantes se encontraron en el caso del factor de estabilidad emocional, aunque tales resultados no habían sido anticipados en una hipótesis.

Por lo que se refiere a la Hipótesis 3, los resultados han mostrado que los efectos del faking sobre la validez de las medidas de personalidad EF quasi-ipsativas son independientes de que el criterio se evalúe mediante escalas de valoración (juicios) o con datos objetivos como la CAP, es decir en ambos casos se produce una reducción en la correlación observada. No obstante, las correlaciones tanto observadas como corregidas muestran que la magnitud de la reducción es mayor en el caso de la evaluación del desempeño mediante escalas de valoración.

Los resultados de la Tabla 52 deben ponerse también en relación con las Hipótesis 4 y 5 de este estudio. Si se observan los coeficientes de fiabilidad en condiciones honestas y faking y el valor de la restricción del rango puede comprobarse que las hipótesis derivadas de la teoría del faking han sido apoyadas empíricamente, por lo que tales resultados dan un apoyo importante a la teoría psicométrica del faking, añadiendo nueva evidencia en apoyo de la misma en relación con un tipo de cuestionario de personalidad que no había sido investigado hasta el presente.

El último análisis que se presenta en esta tesis tiene que ver con la predicción del criterio de desempeño más amplio posible con las medidas realizadas en esta tesis. Hemos combinado en un único compuesto las cuatro medidas de desempeño realizadas, para lo cual se ha realizado un análisis de componentes principales y se han salvado las puntuaciones factoriales de cada sujeto. Seguidamente, tales cargas se emplearon como variable para correlacionar con las puntuaciones de los Big Five tanto en respuesta honesta como en respuesta faking. El análisis de componentes principales mostró que sólo había un componente significativo, con una raíz latente de 2.123 (el siguiente componente tenía una raíz de .860) y que explicaba el 53.1% de la varianza. Las cargas factoriales fueron .502, .859, .750 y .756 para CPA, CDTE, CDCE y CDAN, respectivamente. Por tanto, un compuesto global del desempeño académico explica adecuadamente la relación entre las cuatro medidas.

En la Tabla 54 se presentan las correlaciones observadas y corregidas entre los cinco factores y el compuesto de desempeño académico total, obtenidas tanto para la condición honesta como para la condición faking.

Tabla 54. Correlaciones entre los Big Five en Condiciones Honesta y Faking con el Desempeño Académico Total Empleando la Muestra del Diseño Intra-Grupo

Criterio	EE <sub>H</sub>	EX <sub>H</sub>	AE <sub>H</sub>	AM <sub>H</sub>	CO <sub>H</sub>	EE <sub>F</sub>	EX <sub>F</sub>	AE <sub>F</sub>	AM <sub>F</sub>	CO <sub>F</sub>
DAT – Obs	-.11	-.04	-.09	-.04	.37	-.11	-.03	.06	-.02	.17
<b>DAT – Corr</b>	<b>-.16</b>	<b>-.06</b>	<b>-.12</b>	<b>-.06</b>	<b>.50</b>	<b>-.21</b>	<b>-.05</b>	<b>.10</b>	<b>-.03</b>	<b>.24</b>

Como puede observarse la correlación corregida de conciencia con el compuesto de desempeño total en la condición honesta es de .50, es decir, un 6.4% mayor que la

correlación obtenida para el desempeño global evaluado por las escalas de valoración. En el caso de la condición faking, la correlación corregida ha sido de .24, es decir un 20% mayor que la correlación para el desempeño global. Estos dos resultados, tomados en conjunto, indican que el factor de conciencia predice aún mejor un compuesto amplio de desempeño académico que incluya tanto evaluación de juicios como datos de resultados.

## DISCUSIÓN

Este estudio tenía tres objetivos principales. Primero determinar si las medidas de personalidad EF quasi-ipsativas predicen el desempeño y los resultados académicos tanto en condiciones honestas como de faking y, particularmente, examinar la validez predictiva del factor de conciencia. Segundo, comprobar si el método de medida del desempeño se asocia con un tamaño diferente de validez predictiva. A este respecto se planteó la hipótesis de que el factor de conciencia predice mejor el desempeño medido con escalas de valoración que con datos de resultados académico. El tercer objetivo era comprobar si los efectos del faking (p. ej., reducción del tamaño de validez) se producen con independencia del método de medición del desempeño (juicios o resultados).

En relación con el primer objetivo, se examinó, en primer lugar, la validez predictiva de la personalidad en condiciones honestas y se presentaron los resultados de la muestra total y de dos submuestras (a) una submuestra inter-grupos y (b) una submuestra intra-grupo. Los resultados obtenidos para la condición de respuesta honesta mostraron que el factor conciencia es el mejor predictor del desempeño cuando se utilizan medidas de EF quasi-ipsativas con valores de  $r$  significativos en todos los casos, siendo las correlaciones más altas en el diseño inter-grupos donde, a excepción de la relación con el desempeño contextual ( $r = .15$ ), el valor de  $r$  fue superior a .30. El mejor estimador de la validez del factor de conciencia para predecir el desempeño académico total de los estudiantes es de .50 (ver Tabla 54) en la condición honesta y de .24 en la condición de faking. Estos resultados coinciden con los obtenidos por Salgado y Táuriz (2014) y Salgado et al. (2015) quienes encontraron que el factor conciencia era el mejor predictor del desempeño académico. Esta es, por tanto, la primera contribución única de este estudio.

También se pudo observar que estabilidad emocional y extraversión eran predictores de varios criterios de desempeño académico en la condición honesta. No obstante, las correlaciones significativas que se obtienen entre estos factores y las medidas

de desempeño varían en cada muestra analizada. Así, por ejemplo, el factor extraversión muestra correlaciones significativas con todos los criterios de desempeño evaluados a excepción de la medida de desempeño de tarea en la muestra inter-grupos, mientras que no se encontraron correlaciones significativas con este factor en la muestra intra-grupo. Con los factores apertura a la experiencia y amigabilidad ocurrió una situación similar, y aunque no han destacado como predictores del desempeño, académico, han obtenido correlaciones significativas con alguno de los criterios de desempeño analizados en el diseño intra-grupo (en el caso del factor amigabilidad también en la muestra total). Estos resultados ponen de manifiesto la relevante variabilidad que se produce entre los diseños experimentales de la condición honesta.

En relación con el primer objetivo, en segundo lugar, se examinaron los resultados obtenidos en la condición faking en un diseño intra-grupo. Los resultados ponen de manifiesto la robustez del factor conciencia como predictor del desempeño académico. Aun bajo condiciones de faking, este factor ha mostrado correlaciones significativas con casi todas las variables de desempeño analizadas, a excepción del criterio de desempeño contextual, con tamaños observados del efecto aceptables entre .10 y .15, que se aproximan a los obtenidos para medidas tradicionales en condiciones honestas (Salgado et al., 2015). Esta es, por consiguiente, la segunda contribución única de este estudio. Asimismo, se han encontrado correlaciones significativas con el factor estabilidad emocional para los criterios de desempeño de tarea y contextual, lo que muestra que este factor es, tras conciencia, un predictor válido del desempeño incluso en condiciones de faking. Estos resultados siguen la línea de los obtenidos en la condición honesta de este estudio y de los resultados de meta-análisis previos sobre la validez predictiva de las medidas de EF quasi-ipsativas (en condiciones honestas) que indicaron que, tras conciencia, estabilidad emocional, es un predictor adecuado del desempeño (Salgado et al., 2015). Finalmente, los factores extraversión, apertura a la experiencia y amigabilidad no han destacado como predictores del desempeño en esta condición.

Los resultados obtenidos muestran, por tanto, que la personalidad evaluada con medidas de EF quasi-ipsativas predice el desempeño incluso en condiciones de faking. No obstante, existe una reducción de los tamaños del efecto de las correlaciones en esta condición, aunque la misma no puede ser totalmente atribuible al faking en el presente estudio, ya que cuando se examinan las correlaciones corregidas por error de medida y restricción en el rango no se aproximan a los valores de las correlaciones en condiciones de respuesta honesta.

Como hemos indicado en otro lugar, los resultados de las muestras honestas reflejan una importante variabilidad en los resultados entre los diseños experimentales lo que muestra que existen otras variables, no el faking, que podrían estar afectando a los resultados.

Con respecto al segundo objetivo, relativo a si la validez predictiva es diferente según cual sea el método de medida del desempeño, este estudio ha contribuido mostrando que el factor conciencia predice mejor el desempeño medido con escalas de valoración que con datos de resultados académico. Esta contribución también es única de este estudio y tesis, ya que no ha habido investigación previa que haya examinado esta cuestión. Por otro lado, los resultados se unen a la creciente evidencia empírica que indica que el método de medida del desempeño es un potente moderador de la validez de los instrumentos predictivos, por ejemplo, de los tests de capacidad cognitiva (Salgado y Moscoso, 2019) y la entrevista de selección (McDaniel et al., 1994). La evidencia aportada indica que también en el caso de la medida de la personalidad el método de medición del desempeño puede tener importantes efectos sobre la validez, hasta tal punto que el tamaño de los coeficientes de validez puede llegar a ser más del doble, como se muestra en la Tabla 53, donde el coeficiente corregido para las medidas de resultados es de .21 y el coeficiente para las escalas de valoración es de .47; es decir un 124% mayor en el segundo caso.

En relación con el tercer objetivo referente a comprobar si los efectos del faking (p. ej., reducción del tamaño de validez) se producen con independencia del método de medición del desempeño (juicios o resultados), este estudio ha contribuido mostrando que el efecto de faking se produce tanto cuando se emplean datos de juicios (escalas de valoración) como datos de resultados (calificación académica promedio), aunque la reducción en el coeficiente de validez fue mayor cuando el desempeño se midió con datos de juicios.

Una contribución adicional de este estudio tiene que ver con la teoría psicométrica del faking (Salgado, 2016). De acuerdo a esta teoría, cuando se instruye a los sujetos a falsear las respuestas a los cuestionarios de personalidad o cuando los sujetos están motivados por otras razones (p. ej., selección de personal o procesos de admisión universitarios), los sujetos pueden falsear dichos cuestionarios y, como consecuencia, la fiabilidad y la validez de los cuestionarios se verá atenuada, debido a un incremento del error de medida y una reducción en el rango de las puntuaciones. Este efecto ha podido comprobarse en el presente estudio como quedó reflejado en la Tabla 52. En dicha tabla



puede verse que el coeficiente alfa de consistencia interna de las medidas de los factores de personalidad en el caso de la respuesta bajo instrucciones de faking era menor que cuando fueron contestadas en condiciones de respuesta honesta. También pudo observarse que se producía cierto grado de restricción en el rango en cuatro de los factores de personalidad (la excepción fue el factor de conciencia). Por tanto, el estudio ha contribuido de modo único poniendo a prueba las predicciones de la teoría psicométrica del faking en unas condiciones y con un tipo de cuestionario de personalidad no examinado hasta ahora. Los resultados proporcionan evidencia empírica de apoyo a las predicciones de la teoría.

#### IMPLICACIONES DE LOS RESULTADOS PARA LA TEORÍA Y LA PRÁCTICA DE LA EVALUACIÓN DE LA PERSONALIDAD

Los resultados de este estudio tienen implicaciones tanto para la teoría como para la práctica de la evaluación de la personalidad en contextos aplicados.

En relación con la teoría de la evaluación de la personalidad, una implicación relevante de los resultados tiene que ver con el hecho de que los efectos psicométricos del faking no parecen ser los únicos factores relevantes o determinantes en la reducción de la validez predictiva de las medidas de personalidad. Cuando los efectos derivados del faking sobre la fiabilidad (consistencia interna) y la restricción en el rango han sido controlados y los coeficientes de validez han sido corregidos, todavía ha podido observarse una notable diferencia entre los coeficientes de validez obtenidos en condiciones de respuesta honesta y de respuesta con faking, que no debería haber sido tan grande una vez que los efectos habían sido corregidos psicométricamente.

Por ello, especulamos que el cambio (reducción) en los coeficientes de validez predictiva puede deberse también a otros factores idiosincráticos (p. ej., cambios en el modo de respuesta de los individuos) o de cansancio (respuesta a un cuestionario grande en dos ocasiones seguidas, que requerían más de una hora de trabajo) o de práctica (menor implicación en la segunda ocasión, con respuestas menos elaboradas). Futuros estudios deberían examinar la potencial contribución de estos factores (y otros) a la reducción de la validez predictiva.

En relación con la práctica de la evaluación de la personalidad, una implicación importante es que el tamaño de validez predictiva es considerable cuando se emplea un cuestionario de personalidad de EF quasi-ipsativa y que la predicción aun es mejor cuando



se examinan criterios amplios del desempeño académico. Por ello, una recomendación para los profesionales de la evaluación en contextos aplicados (p. ej., procesos de admisión de estudiantes, procesos de selección para internship y formación) es emplear cuestionarios de EF quasi-ipsativos ya que además de buenos predictores son robustos contra faking. Además, recomendamos que los coeficientes de validez observados se corrijan por restricción en el rango para establecer un estimador de la validez menos sesgado. Por último, los profesionales deben ser conscientes de que la validez de los cuestionarios de personalidad (en nuestro caso, los cuestionarios de EF quasi-ipsativos) no es idéntica para todos los modos de medición del desempeño académico y deben emplear el coeficiente apropiado para el tipo de medida del criterio que se va a emplear en cada caso, recordando que los coeficientes de validez son menores para los datos de resultados que para los datos de juicios.

#### LIMITACIONES DEL ESTUDIO

Este estudio, como todas las investigaciones empíricas, no está exento de limitaciones. Una primera limitación proviene de las diferencias en los tamaños de las muestras, lo que produce diferentes errores de muestreo. Asimismo, las características de los sujetos también pueden afectar a los resultados. Recordemos que provienen de diversas titulaciones y esto puede afectar de modo muy específico a criterios como el CAP lo que puede estar condicionando los resultados en la condición honesta y también afectar los resultados de la condición faking.

Una segunda limitación se deriva de que en la condición faking sólo se han analizado los resultados en un diseño intra-grupo, no habiéndose podido obtener un número suficiente de sujetos en este estudio para poder examinar los resultados en un diseño inter-grupos. Se espera poder continuar con esta investigación y obtener datos suficientes que permitan comprobar si estos resultados se mantienen también en dicho diseño.

También cabe resaltar que la condición faking de este estudio se trata de una condición de máxima distorsión en la que se induce a los participantes a cometer un elevado grado de faking. Es decir, los resultados se han obtenido en una condición de desempeño máximo en el cuestionario de personalidad y, por lo tanto, se espera que los resultados del faking sean menos intensos en condiciones de desempeño normal. En el presente caso no se ha podido controlar este variable (faking en desempeño máximo vs.

faking en desempeño típico), lo que es una tercera limitación del estudio. Por ello sería recomendable llevar a cabo estudios que analicen la validez de criterio de los cuestionarios de personalidad de EF quasi-ipsativos en contextos reales de selección para examinar si se reproducen los resultados obtenidos en este estudio.

Asimismo, en este estudio se ha empleado una medida de EF quasi-ipsativa para predecir criterios de desempeño académicos exclusivamente, lo que no permite la generalización de los resultados a otro tipo de cuestionarios de EF y otros criterios. Recientes estudios (p. ej., Fisher et al., 2019; Salgado, 2018) han mostrado que incluso entre cuestionarios de EF quasi-ipsativos, el método de obtención de la puntuación en los Big Five (basados en la teoría clásica de la medida, con o sin dependencia algebraica o en los modelos thurstonianos de teoría de respuesta al ítem) pueden producir coeficientes de validez predictiva notablemente distintos. En ese sentido, en futuras investigaciones resultaría de interés examinar si el faking afecta a la validez predictiva de otros tipos de cuestionarios de EF (ipsativos y normativos) y sobre otros criterios como, por ejemplo, el desempeño laboral.

No obstante, las limitaciones señaladas, este estudio ha realizado varias contribuciones únicas sobre los efectos del faking en la validez predictiva de las medidas de personalidad de EF quasi-ipsativas y, como se ha mencionado, los resultados han mostrado que, incluso en condiciones de faking, la personalidad medida con cuestionarios EF quasi-ipsativos predice el desempeño académico de los estudiantes.



## CONCLUSIONES GENERALES





## CONCLUSIONES GENERALES

El faking es uno de los principales problemas de la evaluación de la personalidad, especialmente en los contextos de selección personal y en los procesos de admisión de alumnos en las universidades. Las consecuencias negativas del faking sobre las puntuaciones, las propiedades psicométricas y la potencial alteración del ranking de los evaluados cuando se emplean instrumentos de evaluación de la personalidad *single-stimulus* (p. ej., cuestionarios con formatos de respuesta tipo Likert), ha llevado a los investigadores y profesionales de la evaluación a la búsqueda de herramientas alternativas que controlen este fenómeno. Entre los procedimientos evaluativos alternativos se encuentran los cuestionarios de elección forzosa.

Los cuestionarios de EF son un método de evaluación de la personalidad que podría reducir los efectos del faking, al caracterizarse por obligar a las personas a realizar una elección entre ítems de similar grado de deseabilidad social, reduciendo, de este modo, las posibilidades de responder escogiendo la opción más deseable o más favorable para los intereses de quien responde al cuestionario. No obstante, aunque se habla de los cuestionarios de EF como una entidad única, lo cierto es que existen diferencias entre ellos debido a la naturaleza de las puntuaciones que pueden obtenerse y de las características de dichas puntuaciones (Cattell, 1948; Clemans, 1966; Hicks, 1970; Salgado y Táuriz, 2014). Así, por lo general, se considera que existen tres tipos de cuestionarios de EF en función de la naturaleza de las puntuaciones que proporcionan (a) normativos, (b) ipsativos y (c) quasi-ipsativos.

A pesar de ser instrumentos que presentan a priori ventajas metodológicas relevantes en relación con el control del faking, todavía falta investigación sobre cuestiones tales como su grado de resistencia al faking y los posibles efectos del faking sobre las propiedades de la medida, cuestiones que o bien no han sido investigadas en su totalidad o bien no lo han sido suficientemente para lograr una respuesta concluyente o que haya suscitado un elevado consenso.

Esta tesis doctoral, que se ha realizado teniendo como marco teórico el modelo psicométrico del efectos del faking (Salgado, 2016), ha tenido como propósito principal investigar tres aspectos específicos de los efectos del faking sobre los cuestionarios de EF que no habían sido examinados previamente. El primer objetivo ha sido estudiar la resistencia de los tres tipos de cuestionarios de EF (es decir, normativos, ipsativos y quasi-ipsativos) a los efectos del faking sobre las distribuciones de las puntuaciones (media y SD) y la fiabilidad, con la finalidad de comprobar si existen diferencias relevantes entre las condiciones de respuesta honesta y faking cuando se utilizan los cuestionarios de EF. El segundo objetivo de esta investigación ha sido analizar los efectos del faking sobre la invarianza o equivalencia de la medida, es decir, sobre la equivalencia de la estructura factorial de los cuestionarios de EF quasi-ipsativos para comprobar si dicha estructura factorial se mantiene invariante o, por el contrario, se altera en condiciones de faking. Por último, como tercer objetivo se ha examinado si el faking afecta a la validez de criterio de los cuestionarios de EF quasi-ipsativos, y a la validez del factor conciencia en particular, para predecir el desempeño y los resultados académicos tanto en condiciones honestas como de faking. Para dar respuestas a estas estas cuestiones se realizaron tres estudios empíricos cuyas conclusiones se presentan a continuación.

El meta-análisis presentado en el Estudio 1 es único en el sentido de que incluye un meta-análisis jerárquico de moderadores (tipo de diseño y formato del cuestionario) lo que ha permitido obtener resultados más precisos sobre las cuestiones analizadas y que no había sido realizada hasta el momento. Por ello, ha supuesto un avance significativo en el conocimiento de la resistencia de los cuestionarios de EF al faking.

En primer lugar, los resultados del meta-análisis han indicado que los cuestionarios de EF (en los tres formatos analizados) tienen resistencia al faking. De hecho, si se comparan los tamaños del efecto obtenidos en este estudio con los hallados por Viswesvaran y Ones (1999) para los cuestionarios de personalidad SS, se puede comprobar que, en general, los tamaños del efecto proporcionados por nuestro meta-análisis son más pequeños que los obtenidos por Viswesvaran y Ones (1999), lo que indica que los cuestionarios EF presentan más resistencia al faking que las medidas SS.

Con relación a los meta-análisis jerárquicos, los resultados obtenidos permiten afirmar que los dos moderadores, el tipo de cuestionario de EF (normativo, ipsativo o quasi-ipsativo) y el tipo de diseño experimental, afectan a los resultados.

Se ha encontrado que el tipo quasi-ipsativo es más resistente al faking que los otros tipos de cuestionarios, especialmente en el factor de conciencia que es el más



relevante desde el punto de vista predictivo del desempeño ocupacional y del desempeño académico.

Con respecto al tipo de diseño, los hallazgos indican que el faking se produce en mayor grado en los estudios experimentales, particularmente en los que emplean diseños intra-grupo, mientras que los diseños correlacionales con muestras de candidatos y empleados presentan tamaños del efecto notablemente más reducidos o irrelevantes, siendo los tamaños más pequeños los encontrados con empleados, en los que solo se aprecia faking en dos factores, extraversión y apertura, y en ambos casos de escasa magnitud. No obstante debe precisarse que el diseño típico antes-después empleado en el estudio del faking no es lo suficientemente potente para descartar los efectos de los errores de medida transitorios (*transient error*), lo que sugiere que una porción del tamaño del efecto atribuido al faking en tales diseños podría ser una consecuencia del error transitorio. Por todo lo anterior, se puede afirmar que la magnitud del faking puede ser en parte un fenómeno de laboratorio. Por un lado, los tamaños de efecto más altos encontrados en los estudios de laboratorio no se producen en los escenarios reales de selección y con empleados. Por otro lado, los diseños experimentales utilizados no controlaron el efecto del error de medida transitorio, cuyos efectos se asignan al faking.

El Estudio 2 es también una aportación única de esta tesis al conocimiento del faking y sus efectos en los cuestionarios de EF, ya que se trata del primer estudio que ha examinado la invarianza o equivalencia de la medida de un test de EF quasi-ipsativo en condiciones de faking. Este estudio es relevante también por el hecho de haber examinado la invarianza de la medida en tres muestras y con dos diseños experimentales (inter e intra-grupo), lo que ha permitido determinar el posible efecto moderador del diseño en los resultados. Además, el examen no se limitó a la realización de un análisis estructural único, sino que se llevaron a cabo tres tipos de análisis de la estructura latente. Se realizaron análisis factoriales exploratorios (método de máxima verosimilitud) análisis de componentes principales y análisis factorial confirmatorio. En el caso de los dos tipos de análisis factoriales se calcularon los índices de bondad de ajuste y de test de equivalencia y en el caso de los análisis de componentes principales se calcularon los Coeficientes de Burt-Tacker, para obtener un índice más robusto de las similitudes entre las estructuras factoriales de ambas condiciones. Se trata por tanto del estudio más exhaustivo realizado hasta el momento con cuestionarios de EF quasi-ipsativos sobre esta cuestión. Los resultados hallados mostraron: (a) el buen ajuste de una estructura de cinco factores o componentes a los datos en los dos tipos de análisis exploratorios; (b) índices de bondad

de ajuste buenos o excelentes en los dos tipos de análisis factoriales y (c) coeficientes de congruencia entre componentes convergentes muy altos en todas las condiciones analizadas.

Por tanto, estos resultados han demostrado la robustez de la estructura factorial de cinco factores en los cuestionarios de EF quasi-ipsativos que resultó invariante tanto en condiciones de respuesta honesta como bajo instrucciones de faking. Resultados que apoyan la validez de constructo de este tipo de cuestionarios basados en el modelo de los Big Five para evaluar la personalidad basada en este modelo y, por tanto, los convierten una herramienta apropiada para su uso en contextos aplicados de selección.

El Estudio 3 también supone una aportación empírica única ya que es el primero que ha examinado simultáneamente los efectos que la validez de criterio de los cuestionarios de EF quasi-ipsativos tiene el hecho de contestar honestamente y en condiciones de faking. Se ha examinado la validez predictiva de estas medidas de personalidad para criterios académicos y se han comparado los resultados entre un diseño inter-grupos y un diseño intra-grupo.

Los resultados permiten concluir que la personalidad evaluada con cuestionarios de EF quasi-ipsativos predice el desempeño incluso en condiciones de faking. Se ha constatado que el factor conciencia es el mejor predictor del desempeño académico, independientemente de la condición (honesto o faking) o el diseño experimental en el que se evalúe. Estos resultados concuerdan con los obtenidos en el meta-análisis de Salgado y Táuriz (2014) sobre la validez predictiva de las medidas de EF.

El Estudio 3 ha permitido también analizar el efecto moderador que tiene el tipo de medida de desempeño sobre la validez predictiva de los cuestionarios de EF quasi-ipsativos en condición honesta y de faking. Se encontró que el factor conciencia predice mejor el desempeño medido con escalas de valoración que con datos de resultados (calificaciones) en ambas condiciones. Estudios previos han mostrado, por ejemplo, que las capacidades cognitivas presentan coeficientes de validez diferentes para predecir el desempeño en el trabajo, según se haya medido éste con escalas de valoración o con datos de producción (p. ej., Salgado y Moscoso, 2019). Sin embargo, esta cuestión no había sido analizada en el ámbito del desempeño académico. Los resultados, han demostrado en nuestro estudio que el tipo de medida de desempeño es un potente moderador de la validez de los instrumentos predictivos. Respecto a esta cuestión también se ha demostrado que el faking reduce los coeficientes de validez tanto si se emplean datos de

juicios como datos de resultados, aunque se ha observado una mayor reducción de la validez cuando se usan datos de juicios.

Finalmente, el Estudio 3 ha proporcionado evidencia empírica que apoya las predicciones de la teoría de psicometría del faking (Salgado, 2016). Los resultados han demostrado que los coeficientes de fiabilidad son más pequeños y se produce cierto grado de restricción en el rango (con la excepción de factor conciencia) en la condición faking con respecto a la honesta. Siendo el primer estudio que analiza este efecto en una medida de EF quasi-ipsativa.

#### IMPLICACIONES PARA LA PRÁCTICA

Los resultados obtenidos en los tres estudios realizados tienen implicaciones relevantes desde el punto de vista práctico en lo que a los procesos de selección y evaluación se refiere.

Una primera implicación es que los cuestionarios de EF quasi-ipsativos han demostrado ser una medida de personalidad robusta contra el faking. Los resultados han puesto de manifiesto que es una medida resistente al faking, con un elevado grado de equivalencia de medida en condiciones de faking y honestas y, además, ha mostrado ser un adecuado predictor del desempeño incluso en condiciones de faking. Por todo ello, podemos afirmar que los cuestionarios de EF en formato quasi-ipsativo son una herramienta útil para el control del faking y válida para la evaluación de la personalidad. Así, se proponen como una alternativa adecuada a las medidas SS para la evaluación de la personalidad en contextos aplicados que puedan motivar faking.

Una segunda implicación para la evaluación de personalidad en contextos académicos aplicados es que los cuestionarios de EF quasi-ipsativos han mostrado una validez predictiva importante sobre diferentes criterios del desempeño académico. Por lo tanto, sería muy recomendable que se emplee este tipo de medidas en los procesos de evaluación (procesos de admisión de estudiantes, procesos de selección para formación) por la doble razón de ser buenos predictores del desempeño académico y de ser resistentes al faking.

Una tercera implicación práctica de los resultados de esta tesis es que se ha mostrado que la validez de los cuestionarios de personalidad varía en función de la medida que se utilice para evaluar el desempeño académico. Por lo tanto, los evaluadores

deben tener presente qué medida de desempeño pretenden predecir, ya que no se predicen por igual medidas distintas (juicios y datos de resultados) del desempeño académico.

#### SUGERENCIAS PARA INVESTIGACIONES FUTURAS

Esta tesis ha supuesto una contribución al conocimiento de los efectos del faking sobre las medidas de personalidad de EF y ha puesto de manifiesto la utilidad de las mismas para el control del faking. No obstante, existen aspectos que es necesario seguir investigando en este ámbito. En ese sentido se proponen algunas sugerencias para futuras investigaciones.

En el Estudio 1 se pudo observar que en varios casos se ha producido error de muestreo de segundo orden y que, además, no se ha podido analizar la resistencia al faking en los cuestionarios de EF normativos, en ambos casos por el escaso número de estudios existente. Por ello se recomienda seguir realizando estudios primarios que examinen la capacidad de los cuestionarios de EF, en todos sus formatos, para reducir el faking con el objetivo de ampliar el meta-análisis presentado.

También es necesario seguir investigando los efectos del faking sobre la validez de constructo y de criterio para ampliar el conocimiento sobre esta cuestión. Los resultados presentados en esta tesis apoyan las predicciones del modelo psicométrico del faking, pero estudios adicionales son necesarios. En ese sentido, la segunda sugerencia de investigación que se plantea es realizar réplicas de estos estudios con nuevas muestras en el contexto experimental pero también ampliar su estudio a un contexto real (candidatos vs empleados) para obtener una mayor evidencia y apoyo empírico a los resultados encontrados.

También cabe señalar, la necesidad de ampliar las investigaciones aquí realizadas en contextos de faking a otros formatos de cuestionarios de EF (ipsativos y normativos) así como a otros criterios (desempeño laboral, satisfacción, etc.) con el objetivo de acumular estudios primarios para poder llevar a cabo meta-análisis que presenten resultados más robustos sobre la relación entre estas variables.

#### LIMITACIONES

Como toda investigación empírica esta tesis doctoral también presenta limitaciones. La primera de ellas es que no se ha podido realizar un meta-análisis

jerárquico para el formato normativo, lo que ha impedido comparar la resistencia al faking entre los tres tipos de puntuaciones de EF (normativas, ipsativas y quasi-ipsativas). Asimismo, en este análisis se han incluido medidas de EF tanto basadas en los Big Five como en otros modelos de personalidad, por lo que sería recomendable analizar el efecto que tiene el faking sobre estos modelos por separado para comprobar si se corroboran los resultados encontrados.

Por lo que respecta a los Estudios 2 y 3, se han realizado en un contexto experimental de laboratorio con una muestra muy específica, estudiantes universitarios, por lo que los resultados obtenidos pueden estar afectados por las características particulares de estas muestras. Otro punto a tener en cuenta es que los análisis no se han llevado a cabo con muestras totalmente independientes, dado que la muestra intra-grupo formaba parte de la muestra total analizada.

Asimismo, los datos de la condición faking sólo se han obtenido bajo un contexto de desempeño máximo, es decir, se ha inducido a los sujetos, con instrucciones específicas, a cometer faking, por lo que falta información sobre qué resultado se obtendrían en condiciones de desempeño normal. Aunque se espera que los tamaños el efecto sean más bajos, no se ha podido realizar esta comparación (faking en desempeño máximo vs faking en desempeño típico).

Una última limitación de estas investigaciones es que estos estudios se han llevado a cabo con una medida de EF quasi-ipsativa, por lo que se desconoce de qué manera afectaría el faking a la validez predictiva y de constructo de las medidas ipsativas y normativas.

## SUMARIO

El faking es un problema importante para la evaluación de la personalidad en contextos académicos y laborales. Los cuestionarios de personalidad EF han surgido como una herramienta que presenta ventajas metodológicas para controlar y reducir el faking. Sin embargo, varios los aspectos sobre este tipo medida y su resistencia al faking no habían sido investigados.

Poniendo a prueba las predicciones del modelo psicométrico de efectos del faking, esta tesis ha contribuido al conocimiento de los efectos del faking sobre las medidas de personalidad de EF en general, y sobre las medidas quasi-ipsativas en particular, a través de la realización de tres estudios. En el primero se ha constatado que las medidas de EF

presentan resistencia a los efectos del faking. En el segundo estudio se ha podido demostrar que la estructura factorial de cinco componentes de los cuestionarios de EF quasi-ipsativos no se ve afectada por el faking. Y el tercer estudio ha puesto de manifiesto que la personalidad evaluada con este formato de cuestionario predice el desempeño académico incluso en condiciones de faking.

Por todo ello, se puede afirmar que los cuestionarios de EF quasi-ipsativos son una herramienta válida en los procesos de selección y evaluación en contextos aplicados, incluso en condiciones de faking. No obstante, existen otros aspectos sobre los que es necesario seguir profundizando en investigaciones futuras.



## ANEXOS







*Anexo 1. Características y valores de los estudios primarios incluidos en el meta-análisis*

Estudio	N	Formato cuestionario	Diseño	Muestra	EE	$\alpha_{ee}$	EX	$\alpha_{ex}$	AE	$\alpha_{ae}$	A	$\alpha_a$	C	$\alpha_c$
Alonso (2011)	215	I	Intra	Estudiantes	0.14	-	0.82	-	-1.11	-	0.63	-	0.90	-
Anderson, Sison y Wester (1984)	36	N	Inter	Estudiantes	-	-	-	-	-	-	-	-	-	-
Antler, Zaretsky y Ritter (1967)	276	Q	Inter	Estudiantes	-	-	0.35	.85	-	-	-	-	-0.31	.87
Bass (1957)	736	Q	Inter	Candidatos	0.17	.82	0.10	.85	-	-	-	-	0.17	.84
Bowen, Martin y Hunt (2002)	294	I	Inter	Estudiantes	0.13	.77	0.45	.77	-0.02	.77	0.34	.77	0.36	.77
Braun (1962a)	19	Q	Intra	Estudiantes	-	-	-	-	0.31	.83	0.85	.87	2.07	.65
Braun (1962b)	61	Q	Intra	Estudiantes	1.10	-	-	-	-	-	-	-	0.71	-
	53	Q	Intra	Estudiantes	1.79	-	-	-	-	-	-	-	1.65	-
	11	Q	Intra	Estudiantes	1.72	-	-	-	-	-	-	-	1.44	-
	39	Q	Intra	Estudiantes	1.73	-	-	-	-	-	-	-	1.17	-
Braun (1963a)	26	Q	Intra	Estudiantes	-	-	0.61	-	-0.62	-	-1.51	-	1.36	-
Braun (1963b)	24	Q	Intra	Estudiantes	-	-	-	-	-0.02	.83	1.12	.87	1.79	.65
	25	Q	Intra	Estudiantes	-	-	-	-	0.43	.83	1.33	.87	1.30	.65
Braun (1963c)	69	Q	Intra	Estudiantes	1.01	.87	0.66	.85	-	-	-	-	1.14	.87
	44	Q	Intra	Estudiantes	1.15	.87	0.44	.85	-	-	-	-	2.09	.87
Braun (1965)	39	I	Intra	Estudiantes	-	-	1.42	-	-0.37	-	-0.64	-	1.47	-
Braun y Alexander (1961)	37	N	Intra	Estudiantes	1.82	-	-	-	-	-	-	-	-	-
Braun y Farrell (1974)	61	Q	Inter	Estudiantes	-	-	-	-	1.07	.83	0.81	.87	1.21	.65
	90	Q	Inter	Estudiantes	0.95	.87	0.82	.85	-	-	-	-	1.45	.87
Braun y Lafaro (1967)	86	Q	Intra	Estudiantes	0.47	-	0.59	-	1.12	-	0.25	-	1.03	-
Christiansen, et al. (2005)	350	Q	Inter	Estudiantes	-	-	0.47	.84	-	-	-	-	0.40	.81
	122	Q	Inter	Estudiantes	-	-	-	-	-	-	-	-	0.24	.76
Converse, et al. (2010)	107	Q	Inter	Estudiantes	0.94	.88	-	-	-	-	-	-	1.28	.66

Estudio	<i>N</i>	Formato cuestionario	Diseño	Muestra	EE	$\alpha_{ce}$	EX	$\alpha_{ex}$	AE	$\alpha_{ac}$	A	$\alpha_a$	C	$\alpha_c$
	113	Q	Inter	Estudiantes	0.71	.88	-	-	-	-	-	-	1.05	.52
	100	Q	Inter	Estudiantes	0.86	.92	-	-	-	-	-	-	0.46	.69
	104	Q	Inter	Estudiantes	0.75	.88	-	-	-	-	-	-	0.51	.71
Davis (2006)	48	Q	Inter	Empleados	0.06	.82	0.40	.85	0.09	.83	0.95	.81	0.35	.64
Dicken (1959)	19	I	Intra	Estudiantes	-0.90	-	-0.35	-	0.91	-	-0.99	-	1.54	-
Dunnette, et al. (1962)	62	Q	Intra	Empleados	-0.73	-	0.58	-	-1.20	-	0.73	-	-0.27	-
	63	Q	Intra	Candidatos	-0.33	-	0.31	-	-0.49	-	0.10	-	-0.03	-
	96	Q	Intra	Candidatos	-0.40	-	0.64	-	-0.20	-	-0.05	-	-0.42	-
	64	Q	Intra	Candidatos	-0.47	-	0.60	-	-0.99	-	0.49	-	-0.40	-
Estudio 2	653	Q	Inter	Estudiantes	0.47	0.66	-0.24	0.74	-0.55	0.7	0.33	0.84	0.62	.75
	490	Q	Intra	Estudiantes	0.32	0.63	-0.20	0.75	-0.35	0.67	0.12	0.82	0.69	.78
Fineman (1975)	40	N	Inter	Candidatos	-	-	-	-	-	-	-	-	0.08	.55
Ford y Gunderson (1970)	232	I	Inter	Candidatos	-0.37	-	-0.04	-	-0.04	-	0.17	-	-0.14	-
	37	I	Inter	Candidatos	-0.15	-	0.06	-	0.17	-	0.00	-	-0.04	-
Furnham y Craig (1987)	36	I	Inter	Estudiantes	0.36	-	-0.29	-	-0.60	-	-0.55	-	0.25	-
Gordon y Stapleton (1956)	88	Q	Intra	Candidatos	0.03	.87	0.16	.85	-	-	-	-	0.27	.84
	121	Q	Intra	Candidatos	0.28	.74	0.00	.85	-	-	-	-	0.50	.68
Hedberg (1962)	59	Q	Intra	Estudiantes	-	-	0.00	-	0.07	-	-0.19	-	0.11	-
Heggestad, et al. (2006)	590	Q	Inter	Estudiantes	0.61	.80	0.33	.78	0.07	.77	0.13	.78	1.20	.82
Helton y Street (1992)	158	I	Inter	Empleados	-0.61	-	0.25	-	-0.26	-	0.25	-	-0.31	-
Hirsh y Peterson (2008)	203	I	Inter	Estudiantes	0.40	-	-0.21	-	-0.84	-	-0.26	-	0.01	-
Jackson, et al. (2000)	106	Q	Intra	Estudiantes	-	-	-	-	-	-	-	-	0.35	.69
Kaess y Witryol (1957)	507	N	Inter	Empleados	-0.14	-	-	-	-	-	-	-	-	-
	264	N	Inter	Estudiantes	-0.23	-	-	-	-	-	-	-	-	-
	361	N	Inter	Estudiantes	-0.37	-	-	-	-	-	-	-	-	-

Estudio	<i>N</i>	Formato cuestionario	Diseño	Muestra	EE	$\alpha_{cc}$	EX	$\alpha_{cx}$	AE	$\alpha_{ac}$	A	$\alpha_a$	C	$\alpha_c$
Kirchner (1962)	166	I	Inter	Candidatos	-0.07	-	0.16	-	-0.09	-	-0.16	-	-0.11	-
	115	I	Inter	Candidatos	-0.06	-	-0.07	-	-0.05	-	-0.10	-	0.02	-
Kirchner, et al. (1960)	362	I	Inter	Candidatos	-0.62	-	0.32	-	-0.16	-	-0.24	-	0.51	-
Krug (1958)	46	Q	Intra	Estudiantes	1.04	-	-	-	-	-	-	-	1.33	-
Larson, et al. (2013)	253	Q	Inter	Estudiantes	-	-	0.36	0.87	1.27	0.73	-0.39	0.67	1.40	.72
Mahar, et al. (2006)	25	I	Intra	Estudiantes	-	-	1.09	0.89	1.22	0.49	2.15	0.77	1.75	.86
	24	I	Intra	Estudiantes	-	-	0.16	0.6	1.94	0.66	1.82	0.66	2.59	.66
	24	I	Intra	Estudiantes	-	-	0.98	0.66	0.17	0.66	0.63	0.66	0.64	.66
	24	I	Intra	Estudiantes	-	-	0.34	0.66	2.77	0.66	2.36	0.66	1.92	.66
	24	I	Intra	Estudiantes	-	-	1.27	0.66	0.16	0.66	0.47	0.66	0.41	.66
Mahar, Cologon y Duck (1995)	44	I	Intra	Estudiantes	-	-	1.55	0.66	-0.04	0.66	0.00	0.66	0.23	.66
	44	I	Intra	Estudiantes	-	-	0.86	0.66	-0.47	0.66	-1.72	0.66	1.87	.66
Mourer (1959)	19	Q	Inter	Empleados	-1.43	0.82	-0.35	0.85	-	-	-	-	-1.12	.84
Norman (1963)	456	Q	Intra	Estudiantes	0.00	-	0.02	-	0.01	-	0.06	-	0.00	-
O'neil, et al. (2016)	185	Q	Inter	Estudiantes	-	-	0.03	0.78	1.13	0.78	-	-	1.19	.78
	176	Q	Inter	Estudiantes	-	-	0.55	0.78	1.11	0.78	-	-	0.98	.78
	312	Q	Inter	Estudiantes	-	-	-0.03	0.78	0.65	0.78	-	-	0.52	.78
Pavlov, et al. (2018)	180	Q	Intra	Estudiantes	0.69	-	0.60	-	-0.14	-	0.63	-	0.68	-
Rusmore (1956)	81	Q	Intra	Estudiantes	-	0.7	0.10	0.85	-	-	-	-	-	-
Salgado y Lado (2018)	126	Q	Inter	Empleados	-0.09	0.71	0.28	0.73	-0.09	0.66	0.01	0.8	-0.09	.80
Schwab (1971)	11	Q	Intra	Estudiantes	-	-	-	-	0.28	-	0.62	-	0.60	-
	22	Q	Inter	Estudiantes	-	-	-	-	0.27	0.83	0.65	0.87	0.62	.65
	20	Q	Intra	Estudiantes	-	-	-	-	0.79	0.83	1.34	0.87	1.12	.65
	40	Q	Inter	Estudiantes	-	-	-	-	0.79	0.83	2.49	0.87	0.57	.65

Estudio	<i>N</i>	Formato cuestionario	Diseño	Muestra	EE	$\alpha_{cc}$	EX	$\alpha_{ex}$	AE	$\alpha_{ac}$	A	$\alpha_a$	C	$\alpha_c$
Sheppard, et al. (1974)	66	I	Inter	Empleados	0.26	-	0.07	-	-0.15	-	0.12	-	-0.35	-
Spector (1957)	212	N	Inter	Estudiantes	-	-	-	-	0.03	-	-	-	-	-
Stollak (1965)	56	I	Inter	Estudiantes	-	-	-0.01	-	0.71	-	-0.45	-	-0.20	-
Underhill, Lords y Chen (2008)	75	Q	Intra	Estudiantes	-	-	0.00	-	0.00	-	-0.01	-	0.00	-
Vasilopoulos, et al. (2006)	167	Q	Inter	Estudiantes	-	-	-	-	-	-	0.19	0.66	0.27	.51

*Nota:* *N* = tamaño muestral; I = formato de respuesta ipsativo; Q = formato de respuesta quasi-ipsativo; N = formato de respuesta normativo; EE = estabilidad emocional;  $\alpha_{cc}$  = coeficiente de fiabilidad de estabilidad emocional; EX = extraversión;  $\alpha_{ex}$  = coeficiente de fiabilidad de extraversión; AE = apertura a la experiencia;  $\alpha_{ac}$  = coeficiente de fiabilidad de apertura a la experiencia; A = amigabilidad;  $\alpha_a$  = coeficiente de fiabilidad de amigabilidad; C = conciencia;  $\alpha_c$  = coeficiente de fiabilidad de conciencia.

Anexo 2. Matriz de Correlaciones Observadas entre las Variables en la Condición Honesta de la Muestra Total (N = 939)

	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11	V12	V13	V14	V15	V16	V17	V18	V19	V20
V1	-																			
V2	.27	-																		
V3	.33	.38	-																	
V4	.37	.28	.33	-																
V5	-.06	.15	-.06	-.22	-															
V6	.01	.16	.04	-.15	.57	-														
V7	-.08	.09	-.01	-.19	.40	.41	-													
V8	-.01	.17	-.06	-.09	.55	.56	.39	-												
V9	-.08	-.04	-.01	-.07	-.13	-.02	-.04	-.14	-											
V10	-.07	.00	-.06	-.06	-.11	-.10	-.08	-.15	.56	-										
V11	-.12	-.05	-.08	-.07	-.07	-.04	-.01	-.11	.54	.60	-									
V12	-.13	-.10	-.02	-.10	-.10	-.05	-.02	-.15	.58	.57	.58	-								
V13	-.09	-.11	.10	.07	-.19	-.16	-.04	-.15	-.13	-.14	-.18	.02	-							
V14	-.19	-.21	-.05	-.08	-.16	-.14	-.01	-.11	-.13	-.18	-.14	-.02	.48	-						
V15	.06	.14	.11	.16	-.16	-.13	-.06	-.10	-.04	-.06	-.07	.01	.27	.23	-					
V16	-.10	.01	.07	.03	-.13	-.12	.01	-.11	-.15	-.12	-.12	-.01	.44	.43	.27	-				
V17	-.03	-.18	-.09	-.01	-.31	-.35	-.25	-.33	-.09	-.08	-.07	-.12	-.07	.01	-.06	-.03	-			
V18	-.08	-.23	-.15	.02	-.27	-.30	-.24	-.29	-.16	-.10	-.08	-.19	-.09	-.03	-.07	-.10	.44	-		
V19	-.07	-.26	-.24	.00	-.22	-.34	-.26	-.22	-.20	-.12	-.06	-.17	-.18	-.04	-.07	-.08	.44	.55	-	
V20	.01	-.17	-.15	.05	-.22	-.28	-.20	-.20	-.14	-.04	-.02	-.07	-.13	-.01	-.04	-.09	.34	.41	.54	-

*Anexo 3. Matriz de Correlaciones Observadas entre las Variables en la Condición Faking de la Muestra Total (N = 692)*

	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11	V12	V13	V14	V15	V16	V17	V18	V19	V20
V1	-																			
V2	.18	-																		
V3	.23	.22	-																	
V4	.19	.24	.18	-																
V5	-.08	-.01	.03	-.15	-															
V6	.06	.03	.07	-.14	.37	-														
V7	-.00	.08	.03	-.07	.25	.30	-													
V8	-.03	.08	.03	-.15	.36	.42	.25	-												
V9	-.13	.00	-.09	.05	-.15	-.11	-.03	-.15	-											
V10	-.08	.04	-.05	.08	-.03	-.10	-.08	-.07	.38	-										
V11	-.10	.04	-.14	.06	-.08	-.10	-.00	-.14	.40	.47	-									
V12	-.16	-.06	-.02	-.13	-.05	-.08	-.10	-.09	.39	.39	.38	-								
V13	-.03	-.21	.05	-.20	.05	.07	-.08	-.00	-.21	-.20	-.29	.05	-							
V14	-.01	-.19	.02	-.17	-.06	-.04	-.04	-.01	-.15	-.15	-.20	.07	.44	-						
V15	.03	.13	.00	.09	-.17	-.10	.11	-.02	.05	-.04	-.00	-.01	-.06	.07	-					
V16	-.09	-.02	.03	-.11	-.03	-.01	-.01	-.04	-.13	-.11	-.14	.06	.34	.30	.13	-				
V17	.00	-.04	-.10	.09	-.20	-.24	-.12	-.20	-.00	-.06	.02	-.12	-.24	-.15	.00	-.09	-			
V18	-.06	-.03	-.11	.22	-.26	-.26	-.09	-.20	.08	.05	.08	-.15	-.33	-.22	.06	-.17	.41	-		
V19	-.05	-.03	-.21	.20	-.23	-.28	-.08	-.21	.06	.04	.10	-.12	-.34	-.19	.08	-.14	.35	.51	-	
V20	-.05	.02	-.14	.20	-.19	-.27	-.13	-.20	.06	.07	.10	-.13	-.33	-.15	.06	-.16	.31	.44	.53	-



Anexo 4. Matriz de Correlaciones Observadas entre las Variables en la Condición Honesta de la Muestra Intra-grupo (N = 490)

	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11	V12	V13	V14	V15	V16	V17	V18	V19	V20
V1	-																			
V2	.25	-																		
V3	.35	.37	-																	
V4	.35	.29	.32	-																
V5	-.10	.13	-.06	-.23	-															
V6	.01	.17	.06	-.14	.54	-														
V7	-.10	.09	-.01	-.16	.45	.42	-													
V8	-.03	.16	-.05	-.09	.54	.54	.37	-												
V9	-.07	-.06	-.02	-.10	-.13	-.01	-.01	-.12	-											
V10	-.07	-.00	-.06	-.09	-.09	-.09	-.05	-.13	.58	-										
V11	-.15	-.10	-.05	-.08	-.09	-.06	.03	-.09	.52	.58	-									
V12	-.13	-.11	-.03	-.14	-.11	-.06	-.01	-.10	.54	.57	.55	-								
V13	-.04	-.11	.07	.07	-.18	-.13	-.07	-.12	-.11	-.12	-.16	.03	-							
V14	-.15	-.19	-.04	-.04	-.12	-.11	-.02	-.08	-.15	-.20	-.15	-.03	.47	-						
V15	.05	.12	.07	.16	-.13	-.08	-.03	-.07	-.08	-.08	-.08	-.00	.21	.21	-					
V16	-.10	.02	.06	.05	-.11	-.13	.02	-.12	-.09	-.10	-.11	.04	.44	.45	.24	-				
V17	-.03	-.22	-.09	.04	-.35	-.36	-.25	-.39	-.06	-.13	-.06	-.14	-.06	.00	-.11	-.01	-			
V18	-.05	-.27	-.14	.03	-.25	-.28	-.22	-.29	-.14	-.10	-.04	-.18	-.10	-.07	-.10	-.12	.47	-		
V19	-.08	-.27	-.24	.02	-.21	-.35	-.26	-.24	-.24	-.16	-.07	-.16	-.17	-.08	-.08	-.08	.47	.56	-	
V20	-.01	-.15	-.15	.06	-.18	-.24	-.17	-.18	-.16	-.06	-.01	-.09	-.18	-.07	-.06	-.09	.36	.44	.56	-

*Anexo 5. Matriz de Correlaciones Observadas entre las Variables en la Condición Faking de la Muestra Intra-grupo (N = 490)*

	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11	V12	V13	V14	V15	V16	V17	V18	V19	V20
V1	-																			
V2	.19	-																		
V3	.27	.25	-																	
V4	.26	.18	.16	-																
V5	-.09	-.02	.03	-.17	-															
V6	.04	.05	.10	-.13	.41	-														
V7	-.03	.06	.06	-.06	.24	.27	-													
V8	-.03	.09	.02	-.14	.41	.45	.24	-												
V9	-.11	.03	-.05	.01	-.15	-.11	-.00	-.14	-											
V10	-.08	.06	-.05	.08	-.03	-.12	-.07	-.07	.38	-										
V11	-.10	.04	-.13	.00	-.05	-.11	.03	-.11	.39	.51	-									
V12	-.18	-.05	-.05	-.22	-.01	-.05	-.06	-.05	.39	.40	.38	-								
V13	-.03	-.21	.06	-.19	.06	.06	-.07	.01	-.20	-.22	-.30	.09	-							
V14	-.03	-.16	.02	-.14	-.06	-.05	-.03	-.04	-.15	-.16	-.22	.09	.48	-						
V15	.04	.09	.03	.09	-.18	-.14	.11	-.01	.06	-.02	.01	-.01	-.06	.08	-					
V16	-.03	.03	.09	-.11	-.01	-.00	-.01	-.06	-.16	-.13	-.15	.05	.35	.35	.12	-				
V17	.03	-.06	-.09	.12	-.22	-.25	-.13	-.21	-.06	-.09	-.00	-.15	-.28	-.19	.04	-.10	-			
V18	-.05	-.03	-.13	.23	-.25	-.24	-.05	-.22	.11	.05	.08	-.21	-.38	-.28	.09	-.21	.43	-		
V19	-.04	-.05	-.24	.19	-.23	-.28	-.09	-.24	.03	.02	.10	-.13	-.38	-.20	.10	-.18	.39	.55	-	
V20	-.04	-.00	-.16	.21	-.19	-.23	-.12	-.19	.00	.07	.09	-.19	-.39	-.23	.08	-.17	.32	.49	.55	-

Anexo 6. Matriz de Correlaciones Observadas entre las Variables en la Condición Honesta Inter-grupos (N = 449)

	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11	V12	V13	V14	V15	V16	V17	V18	V19	V20
V1	-																			
V2	.30	-																		
V3	.32	.40	-																	
V4	.40	.29	.34	-																
V5	-.03	.18	-.04	-.20	-															
V6	.03	.15	.03	-.14	.62	-														
V7	-.04	.11	.00	-.20	.36	.40	-													
V8	.02	.19	-.05	-.08	.55	.60	.41	-												
V9	-.06	-.01	.02	-.04	-.08	-.02	-.04	-.11	-											
V10	-.05	.02	-.04	-.02	-.10	-.09	-.10	-.13	.53	-										
V11	-.11	.01	-.10	-.06	-.07	-.01	-.02	-.09	.56	.64	-									
V12	-.08	-.06	-.01	-.06	-.04	-.03	-.00	-.15	.61	.58	.61	-								
V13	-.12	-.08	.15	.09	-.18	-.18	.00	-.15	-.13	-.13	-.20	.04	-							
V14	-.22	-.20	-.04	-.10	-.18	-.15	.01	-.14	-.08	-.14	-.10	.00	.50	-						
V15	.10	.17	.16	.19	-.16	-.16	-.06	-.12	.01	-.01	-.04	.03	.34	.27	-					
V16	-.09	.02	.08	.04	-.12	-.09	.02	-.08	-.18	-.13	-.11	-.03	.46	.42	.30	-				
V17	-.01	-.12	-.09	-.05	-.25	-.33	-.22	-.26	-.10	-.02	-.06	-.11	-.05	.03	.01	-.04	-			
V19	-.09	-.18	-.14	.02	-.27	-.30	-.25	-.28	-.17	-.08	-.10	-.19	-.06	.02	-.01	-.06	.42	-		
V18	-.04	-.23	-.23	-.00	-.22	-.31	-.23	-.20	-.13	-.05	-.04	-.15	-.17	.01	-.05	-.07	.43	.55	-	
V20	.04	-.18	-.14	.05	-.23	-.30	-.21	-.21	-.12	-.01	-.02	-.06	-.06	.08	-.01	-.06	.34	.39	.53	-

*Anexo 7. Matriz de Correlaciones Observadas entre las Variables en la Condición Faking Inter-grupos (N = 202)*

	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11	V12	V13	V14	V15	V16	V17	V18	V19	V20
V1	-																			
V2	.11	-																		
V3	.12	.13	-																	
V4	.03	.39	.24	-																
V5	-.06	.01	.03	-.11	-															
V6	.08	-.05	-.05	-.18	.24	-														
V7	.04	.12	-.05	-.10	.27	.36	-													
V8	-.05	.03	.05	-.14	.19	.30	.25	-												
V9	-.17	-.04	-.14	.15	-.16	-.11	-.09	-.16	-											
V10	-.05	-.00	-.02	.08	-.04	-.05	-.11	-.09	.38	-										
V11	-.08	.05	-.14	.20	-.16	-.06	-.08	-.22	.43	.37	-									
V12	-.06	-.06	.07	.08	-.15	-.12	-.17	-.18	.38	.38	.37	-								
V13	.02	-.18	.04	-.25	.04	.13	-.11	-.01	-.26	-.16	-.28	-.09	-							
V14	.05	-.26	.04	-.25	-.07	-.01	-.07	.08	-.16	-.14	-.17	.00	.32	-						
V15	-.00	.21	-.07	.11	-.14	-.02	.10	-.05	.03	-.11	-.03	.00	-.04	.04	-					
V16	-.18	-.10	-.09	-.13	-.05	-.01	-.01	.03	-.08	-.09	-.15	.05	.28	.18	.19	-				
V17	-.04	.04	-.10	-.00	-.11	-.18	-.08	-.17	.13	.04	.06	-.08	-.12	-.04	-.10	-.12	-			
V18	-.02	.04	-.02	.16	-.26	-.31	-.17	-.12	-.00	.03	.07	-.03	-.21	-.08	.01	-.12	.30	-		
V19	-.06	.05	-.12	.20	-.22	-.25	-.05	-.12	.12	.07	.09	-.10	-.22	-.17	.04	-.04	.21	.37	-	
V20	-.04	.10	-.06	.16	-.15	-.35	-.12	-.21	.18	.06	.11	-.02	-.18	.04	.05	-.15	.26	.28	.45	-

## REFERENCIAS





Las referencias marcadas con asterisco (\*) indican los estudios que han sido incluidos en el meta-análisis (Estudio 1).

Aamodt, M. G. (2003). How common is resume fraud? *Assessment Council News*, febrero, 6-7. Recuperado de <http://maamodt.asp.radford.edu/Research%20-%20IO/2003-Feb-Resume%20fraud.pdf>

Ackerman, P. L. (1986). Individual differences in information processing: An investigation of intellectual abilities and task performance during practice. *Intelligence*, 10, 101-139. [https://doi.org/10.1016/0160-2896\(86\)90010-3](https://doi.org/10.1016/0160-2896(86)90010-3)

Adair, C. (2014). *Interventions for Addressing Faking on Personality Assessments for Employee Selection: A Meta-Analysis*. Tesis doctoral no publicada. DePaul University, Chicago, IL.

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, 50, 179-211. [https://doi.org/10.1016/0749-5978\(91\)90020-T](https://doi.org/10.1016/0749-5978(91)90020-T)

Alliger, G. M. y Dwight, S. A. (2000). A meta-analytic investigation of the susceptibility of integrity tests to faking and coaching. *Educational and Psychological Measurement*, 60, 59-72. <https://doi.org/10.1177/00131640021970367>

Alliger, G. M., Lilienfeld, S. O. y Mitchell, K. E. (1996). The susceptibility of overt and covert integrity tests to coaching and faking. *Psychological Science*, 7, 32-39. <https://doi.org/10.1111/j.1467-9280.1996.tb00663.x>

Allport, G. W. y Odbert, H. S. (1936). Trait-names: A psycho-lexical study. *Psychological Monographs*, 47, i-171. <http://dx.doi.org/10.1037/h0093360>

\*Alonso, P. (2011). *Efectos de la deseabilidad social en tests normativos e ipsativos*. Trabajo Fin de Máster no publicado. Universidad de Santiago de Compostela, Santiago de Compostela.

Alonso, P. y Táuriz, G. (2010, julio). *Reduciendo la deseabilidad social: test normativos e ipsativos y estrategia de advertencia*. Póster presentado en el VII Congreso Iberoamericano de Psicología. Oviedo, España.

\*Anderson, H. N., Sison, G. y Wester, S. (1984). Intelligence and dissimulation on the personal orientation inventory. *Journal of Clinical Psychology*, 40, 1394-1398. [https://doi.org/10.1002/1097-4679\(198411\)40:6<1394::AID-JCLP2270400620>3.0.CO;2-R](https://doi.org/10.1002/1097-4679(198411)40:6<1394::AID-JCLP2270400620>3.0.CO;2-R)

Anderson, C. D., Warner, J. L. y Spencer, C. C. (1984). Inflation bias in self-assessment examinations: Implications for valid employee selection. *Journal of Applied Psychology*, 69, 574-580. <http://dx.doi.org/10.1037/0021-9010.69.4.574>



- \*Antler, L., Zaretsky, H. H. y Ritter, W. (1967). The practical validity of the Gordon Personal Profile among United States and foreign medical residents. *The Journal of Social Psychology*, 72, 257-263. <https://doi.org/10.1080/00224545.1967.9922323>
- Arthur, W., Glaze, R. M., Villado, A. J. y Taylor, J. E. (2009). Unproctored Internet-based tests of cognitive ability and personality: Magnitude of cheating and response distortion. *Industrial and Organizational Psychology*, 2, 39-45. <https://doi.org/10.1111/j.1754-9434.2008.01105.x>
- Bahri, M. S. (2013). Faking good in personality and emotional intelligent test: self-enhancement among a cohort of medical school applicants. *Education in Medicine Journal*, 5, 60-71. <https://doi.org/10.5959/eimj.v5i2.135>
- Baron, H. (1996). Strengths and limitation of ipsative measurement. *Journal of Occupational and Organizational Psychology*, 69, 49-56. <https://doi.org/10.1111/j.2044-8325.1996.tb00599.x>
- Barrick, M. R. y Mount, M. K. (1991). The big five personality dimensions and job performance: A meta-analysis. *Personnel Psychology*, 44, 1-26. <https://doi.org/10.1111/j.1744-6570.1991.tb00688.x>
- Barrick, M. R. y Mount, M. K. (1996). Effects of impression management and self-deception on the predictive validity of personality constructs. *Journal of Applied Psychology*, 81, 261-272. <http://dx.doi.org/10.1037/0021-9010.81.3.261>
- Barrick, M. R., Mount, M. K. y Judge, T. A. (2001). The FFM personality dimensions and job performance: Meta-analysis of meta-analyses. Personality and Performance at the Beginning of the New Millennium: What Do We Know and Where Do We Go Next? *International Journal of Selection and Assessment*, 9, 9-30. <https://doi.org/10.1111/1468-2389.00160>
- Bartram, D. (1996). The relationship between ipsatized and normative measures of personality. *Journal of Occupational and Organizational Psychology*, 69, 25-39. <https://doi.org/10.1111/j.2044-8325.1996.tb00597.x>
- Bartram, D. (2005). The Great Eight competencies: A criterion-centric approach to validation. *Journal of Applied Psychology*, 90, 1185-1203. <http://dx.doi.org/10.1037/0021-9010.90.6.1185>
- Bartram, D. (2007). Increasing validity with forced-choice criterion measurement formats. *International Journal of Selection and Assessment*, 15, 263-272. <https://doi.org/10.1111/j.1468-2389.2007.00386.x>
- \*Bass, B. M. (1957). Faking by sales applicants of a forced choice personality inventory. *Journal of Applied Psychology*, 41, 403-404. <http://dx.doi.org/10.1037/h0044670>

- Berman, J. A. (1993). Validation of the P.D.I. Employment Inventory in a retail chain. *Journal of Business and Psychology*, 7, 413-419. <https://doi.org/10.1007/BF01013755>
- Bing, M. N., Kluemper, D., Davison, H. K., Taylor, S. y Novicevic, M. (2011). Overclaiming as a measure of faking. *Organizational Behavior and Human Decision Processes*, 116, 148-162. <https://doi.org/10.1016/j.obhdp.2011.05.006>
- Birkeland, S. A., Manson, T. M., Kisamore, J. L., Brannick, M. T. y Smith, M. A. (2006). A meta-analytic investigation of job applicant faking on personality measures. *International Journal of Selection and Assessment*, 14, 317-335. <https://doi.org/10.1111/j.1468-2389.2006.00354.x>
- Bolino, M. C., Kacmar, K. M., Turnley, W. H. y Gilstrap, J. B. (2008). A multi-level review of impression management motives and behaviors. *Journal of Management*, 34, 1080-1109. <https://doi.org/10.1177/0149206308324325>
- Bolino, M. C. y Turnley, W. H. (2003). More than one way to make an impression: Exploring profiles of impression management. *Journal of Management*, 29, 141-160. <https://doi.org/10.1177/014920630302900202>
- Book, A. S., Holden, R. R., Starzyk, K. B., Wasylkiw, L. y Edwards, M. J. (2006). Psychopathic traits and experimentally induced deception in self-report assessment. *Personality and Individual Differences*, 41, 601-608. <https://doi.org/10.1016/j.paid.2006.02.011>
- Borkenau, P. y Ostendorf, F. (1990). Comparing exploratory and confirmatory factor analysis: A study on the 5-factor model of personality. *Personality and Individual Differences*, 11, 515-524. [https://doi.org/10.1016/0191-8869\(90\)90065-Y](https://doi.org/10.1016/0191-8869(90)90065-Y)
- Boss, P., König, C. J. y Melchers, K. G. (2015). Faking good and faking bad among military conscripts. *Human Performance*, 28, 26-39. <https://doi.org/10.1080/08959285.2014.974758>
- Bott, J. P., O'Connell, M. S., Ramakrishnan, M. y Doverspike, D. (2007). Practical limitations in making decisions regarding the distribution of applicant personality test scores based on incumbent data. *Journal of Business and Psychology*, 22, 123-134. <https://doi.org/10.1007/s10869-007-9053-x>
- Bott, J., Snell, A., Dahling, J. y Smith, B. N. (2010). Predicting individual score elevation in an applicant setting: The influence of individual differences and situational perceptions. *Journal of Applied Social Psychology*, 40, 2774-2790. <https://doi.org/10.1111/j.1559-1816.2010.00680.x>
- \*Bowen, C. C., Martin, B. A. y Hunt, S. T. (2002). A comparison of ipsative and normative approaches for ability to control faking in personality questionnaires.

*The International Journal of Organizational Analysis*, 10, 240-259.  
<https://doi.org/10.1108/eb028952>

Boyce, A. S. (2005). *An investigation of faking: Its antecedents and impacts in applicant settings*. (Tesis doctoral) Recuperado de ProQuest. (1428909).

\*Braun, J. R. (1962a). Effects of a top management faking set on the Gordon Personal Inventory. *Psychological Reports*, 10, 611-614.  
<https://doi.org/10.2466/pr0.1962.10.3.611>

\*Braun, J. R. (1962b). Differential susceptibility to faking of various Ghiselli Self-Description Inventory Scales. *Psychological Reports*, 10, 639-641.  
<https://doi.org/10.2466/pr0.1962.10.3.639>

\*Braun, J. R. (1963a). Effects of positive and negative faking sets on the Survey of Interpersonal Values. *Psychological Reports*, 13, 171-173.  
<https://doi.org/10.2466/pr0.1963.13.1.171>

\*Braun, J. R. (1963b). Fakability of the Gordon Personal Inventory: replication and extension. *The Journal of Psychology*, 55, 441-444.  
<https://doi.org/10.1080/00223980.1963.9916638>

\*Braun, J. R. (1965a). Effects of specific instructions to fake on Gordon Personal Profile scores. *Psychological Reports*, 17, 847-850.  
<https://doi.org/10.2466/pr0.1965.17.3.847>

\*Braun, J. R. (1965b). Note on a faking study with the Myers-Briggs Type Indicator. *Psychological Reports*, 17, 924-924.  
<https://doi.org/10.2466/pr0.1965.17.3.924>

\*Braun, J. R. y Alexander, S. (1961). Fakability of the Short Forced-Choice Anxiety Scale. *Psychological Reports*, 9, 118. <https://doi.org/10.2466/pr0.1961.9.1.118>

\*Braun, J. R. y Farrell, R. M. (1974). Re-examination of the fakability of the Gordon Personal Inventory and Profile: A reply to Schwab. *Psychological Reports*, 34, 247-250. <https://doi.org/10.2466/pr0.1974.34.1.247>

\*Braun, J. R. y LaFaro, D. (1967). Effects of a good impression set on the Thorndike Dimensions of Temperament. *Journal of Educational Measurement*, 4, 237-240.

Brunetti, D. G., Schlottmann, R. S., Scott, A. B. y Hollrah, J. L. (1998). Instructed faking and MMPI-2 response latencies: The potential for assessing response validity. *Journal of Clinical Psychology*, 54, 143-153.  
[https://doi.org/10.1002/\(SICI\)1097-4679\(199802\)54:2<143::AID-JCLP3>3.0.CO;2-T](https://doi.org/10.1002/(SICI)1097-4679(199802)54:2<143::AID-JCLP3>3.0.CO;2-T)

- Buchl, A. K. y Melchers, K. G. (2017). Individual difference variables and the occurrence and effectiveness of faking behavior in interviews. *Frontiers in Psychology*, 8, 686 <https://doi.org/10.3389/fpsyg.2017.00686>
- Burkhart, B. R., Gynther, M. D. y Christian, W. L. (1978). Psychological mindedness, intelligence, and item subtlety endorsement patterns on the MMPI. *Journal of Clinical Psychology*, 34, 76-79. [https://doi.org/10.1002/1097-4679\(197801\)34:1<76::AID-JCLP2270340116>3.0.CO;2-E](https://doi.org/10.1002/1097-4679(197801)34:1<76::AID-JCLP2270340116>3.0.CO;2-E)
- Burns, G. N. y Christiansen, N. D. (2006). Sensitive or senseless: On the use of social desirability measures in selection and assessment. En R. L. Griffith y M. H. Peterson (Eds.), *A closer examination of applicant faking behavior* (pp. 113–148). Greenwich, CT: Information Age.
- Burns, G. N. y Christiansen, N. D. (2011). Methods of measuring faking behavior. *Human Performance*, 24, 358-372. <https://doi.org/10.1080/08959285.2011.597473>
- Callahan, P. (2004). *The cheating culture*. Orlando, FL: Harcourt, Inc.
- Cao, M. (2016). Examining the fakability of forced-choice individual differences measures (Tesis doctoral). Psychology Department. University of Illinois at Urbana-Champaign. Recuperado de: <http://hdl.handle.net/2142/93064>
- Cao, M. y Drasgow, F. (2019). Does forcing reduce faking? A meta-analytic review of forced-choice personality measures in high-stakes situations. *Journal of Applied Psychology*. Publicación previa online. <http://dx.doi.org/10.1037/apl0000414>
- Caprara, G. V. y Perugini, M. (1994). Personality described by adjectives: The generalizability of the Big Five to the Italian lexical context. *European Journal of Personality*, 8, 357-369. <https://doi.org/10.1002/per.2410080502>
- Carvajal, C. C. y Gómez, C. X. (2014). Análisis factorial de ítems de respuesta forzada: Una revisión y un ejemplo. *Revista Latinoamericana de Psicología*, 46, 24-34. [https://doi.org/10.1016/S0120-0534\(14\)70003-2](https://doi.org/10.1016/S0120-0534(14)70003-2)
- Cattell, R. B. (1943). The description of personality: Basic traits resolved into clusters. *The Journal of Abnormal and Social Psychology*, 38, 476-506. <http://dx.doi.org/10.1037/h0054116>
- Cattell, R. B. (1946). *Description and Measurement of Personality*. Oxford, England: World Book Company.
- Cattell, R. B. (1947). Confirmation and clarification of primary personality factors. *Psychometrika*, 12, 197-220. <https://doi.org/10.1007/BF02289253>
- Cattell, R. B. (1948). Concepts and methods in the measurement of group syntality. *Psychological Review*, 55, 48-63. <http://dx.doi.org/10.1037/h0055921>

- Cattell, R. B. y Brennan, J. (1994). Finding personality structure when ipsative measurements are the unavoidable basis of the variables. *American Journal of Psychology*, 107, 261-274. <http://dx.doi.org/10.2307/1423040>
- Cattell, R. B., Cattell, A. K. y Cattell, H. E. P. (1993). *Sixteen Personality Factor Questionnaire. Fifth Edition*. Champaign, IL: Institute for Personality and Ability Testing, Inc.
- Cattell, R. B. y Eber, H. W. (1964). *Sixteen Personality Factor Questionnaire (16PF)*. Champaign, IL: Institute for Personality and Ability Testing, Inc.
- Cellar, D. F., Miller, M. L., Doverspike, D. D. y Klawnsky, J. D. (1996). Comparison of factor structures and criterion-related validity coefficients for two measures of personality based on the five factor model. *Journal of Applied Psychology*, 81, 694-704. <http://dx.doi.org/10.1037/0021-9010.81.6.694>
- Chapman, D. y Webster, J. (2006). Toward an integrated model of applicant reactions and job choice. *The International Journal of Human Resource Management*, 17, 1032-1057. <https://doi.org/10.1080/09585190600696572>
- \*Christiansen, N. D., Burns, G. N. y Montgomery, G. E. (2005). Reconsidering forced-choice item formats for applicant personality assessment. *Human Performance*, 18, 267-307. [https://doi.org/10.1207/s15327043hup1803\\_4](https://doi.org/10.1207/s15327043hup1803_4)
- Christiansen, N. D., Edelstein, S. y Fleming, B. (1998, Abril). *Reconsidering forced-choice formats for applicant personality assessment*. Paper presentado en la 13th annual conference of the Society for Industrial and Organizational Psychology, Dallas, TX.
- Christiansen, N. D., Goffin, R. D., Johnston, N. G. y Rothstein, M. G. (1994). Correcting the 16PF for faking: Effects on criterion-related validity and individual hiring decisions. *Personnel Psychology*, 47, 847-860. <https://doi.org/10.1111/j.1744-6570.1994.tb01581.x>
- Clemans, W. V. (1966). An analytical and empirical examination of some properties of ipsative measures. *Psychometric Monographs*, 14, 1-56.
- Cohen, J. (1977). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ, England: Lawrence Erlbaum Associates, Inc.
- Coleman, N. y Mahaffey, T. (2000). Business student ethics: Selected predictors of attitudes toward cheating. *Teaching Business Ethics*, 4, 121-136. <https://doi.org/10.1023/A:1009855128668>
- Comrey, A. L. (1970). *The Comrey Personality Scales*. San Diego: Educational and Industrial Testing Service.

- Converse, P. D., Oswald, F. L., Imus, A., Hedricks, C., Roy, R. y Butera, H. (2006). Forcing choices in personality measurement. En R. L. Griffith y M. H. Peterson (Eds.), *A closer examination of applicant faking behavior* (pp. 263–282). Greenwich, CT: Information Age.
- \*Converse, P. D., Pathak, J., Quist, J., Merbedone, M., Gotlib, T. y Kostic, E. (2010). Statement desirability ratings in forced-choice personality measure development: implications for reducing score inflation and providing trait-level information. *Human Performance*, 23, 323-342. <https://doi.org/10.1080/08959285.2010.501047>
- Converse, P. D., Peterson, M. H. y Griffith, R. L. (2009). Faking on personality measures: Implications for selection involving multiple predictors. *International Journal of Selection and Assessment*, 17, 47–60. <https://doi.org/10.1111/j.1468-2389.2009.00450.x>
- Cook, T. D. y Campbell, D. T. (1979). The design and conduct of true experiments and quasi-experiments in field settings. En *Reproduced in part in Research in Organizations: Issues and Controversies*. Goodyear Publishing Company.
- Cornwell, J. M. y Dunlap, W. P. (1994). On the questionable soundness of factoring ipsative data: A response to Saville & Willson (1991). *Journal of Occupational and Organizational Psychology*, 67, 89-100. <https://doi.org/10.1111/j.2044-8325.1994.tb00553.x>
- Costa Jr., P. T. y McCrae, R. R. (1987). Neuroticism, somatic complaints, and disease: Is the bark worse than the bite? *Journal of Personality*, 55, 299-316. <https://doi.org/10.1111/j.1467-6494.1987.tb00438.x>
- Costa, P. T. y McCrae, R. R. (1989). *NEO PI/FFI manual supplement for use with the NEO Personality Inventory and the NEO Five-Factor Inventory*. Psychological Assessment Resources.
- Costa Jr., P. T. y McCrae, R. R. (1992a). Four ways five factors are basic. *Personality and Individual Differences*, 13, 653-665. [https://doi.org/10.1016/0191-8869\(92\)90236-I](https://doi.org/10.1016/0191-8869(92)90236-I)
- Costa, P. T. y McCrae, R. R. (1992b). Normal personality assessment in clinical practice: The NEO Personality Inventory. *Psychological Assessment*, 4, 5-13.
- Costa Jr., P. T. y McCrae, R. R. (2008). The Revised NEO Personality Inventory (NEO-PI-R). En G. J. Boyle, G. Matthews y D. H. Saklofske (Eds.), *The SAGE handbook of personality theory and assessment, Vol. 2. Personality measurement and testing* (págs. 179-198). Thousand Oaks, CA, US: Sage Publications, Inc. <http://dx.doi.org/10.4135/9781849200479.n9>



- Costa Jr., P. T., McCrae, R. R. y Kay, G. G (1995). Persons, places, and personality: Career assessment using the Revised NEO Personality Inventory. *Journal of Career Assessment*, 3, 123-139. <https://doi.org/10.1177/106907279500300202>
- Crandall, V. C., Crandall, V. J. y Katkovsky, W. (1965). A children's social desirability questionnaire. *Journal of Consulting Psychology*, 29, 27-36. <http://dx.doi.org/10.1037/h0020966>
- Crowne, D. P. y Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology*, 24, 349-354. <http://dx.doi.org/10.1037/h0047358>
- Damarin, F. y Messick, S. (1965). Response styles as personality variables: a theoretical integration of multivariate research. *ETS Research Bulletin Series*, 1965(1), i-116. <https://doi.org/10.1002/j.2333-8504.1965.tb00967.x>
- Dar-Nimrod, I., Ganesan, A. y MacCann, C. (2018). Coolness as a trait and its relations to the Big Five, self-esteem, social desirability, and action orientation. *Personality and Individual Differences*, 121, 1-6. <http://dx.doi.org/10.1016/j.paid.2017.09.012>
- \*Davis, M. (2006). *The relation between optimism and job performance: An applied setting*. Tesis de máster no publicada. University of Central Florida, Florida, FL.
- \*Dicken, C. F. (1959). Simulated patterns on the Edwards Personal Preference Schedule. *Journal of Applied Psychology*, 43, 372-378. <http://dx.doi.org/10.1037/h0044779>
- Dilchert, S. y Ones, D. S. (2012). Measuring and improving environmental sustainability. En S. E. Jackson, D. S. Ones y S. Dilchert (Eds.), *Managing HR for environmental sustainability*, (págs. 187-221). San Francisco, CA: Jossey-Bass/Wiley.
- Dipboye, R. L. (1992). *Selection interviews: Process perspectives*. South-Western Pub.
- Donovan, J. J., Dwight, S. A. y Hurtz, G. M. (2003). An assessment of the prevalence, severity, and verifiability of entry-level applicant faking using the randomized response technique. *Human Performance*, 16, 81-106. [https://doi.org/10.1207/S15327043HUP1601\\_4](https://doi.org/10.1207/S15327043HUP1601_4)
- Donovan, J. J., Dwight, S. A. y Schneider, D. (2008, abril). *Faking in the real world: Evidence from a field study*. Paper presentado en la 23<sup>rd</sup> annual conference of the Society of Industrial and Organizational Psychology. San Francisco, CA.
- Donovan, J. J., Dwight, S. A. y Schneider, D. (2014). The impact of applicant faking on selection measures, hiring decisions, and employee performance. *Journal of*



- Business and Psychology*, 29, 479-493. <https://doi.org/10.1007/s10869-013-9318-5>
- Douglas, E. F., McDaniel, M. A. y Snell, A. F. (1996). The validity of non-cognitive measures decays when applicants fake. *Academy of Management Proceedings*, 1996, 127-131. <https://doi.org/10.5465/ambpp.1996.4979062>
- Drasgow, F., Stark, S., Chernyshenko, O. S., Nye, C. D., Hulin, C. L. y White, L. A. (2012). *Development of the tailored adaptive personality assessment system (TAPAS) to support Army selection and classification decisions* (Technical Report 1311). Fort Belvoir, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.
- \*Dunnette, M. D., McCartney, J., Carlson, H. C. y Kirchner, W. K. (1962). A study of faking behavior on a forced-choice self-description checklist. *Personnel Psychology*, 15, 13-24. <https://doi.org/10.1111/j.1744-6570.1962.tb01843.x>
- Dwight, S. A. y Donovan, J. J. (2003). Do warnings not to fake reduce faking? *Human Performance*, 16, 1-23. [https://doi.org/10.1207/S15327043HUP1601\\_1](https://doi.org/10.1207/S15327043HUP1601_1)
- Edwards A. L. (1957). *The social desirability variable in personality assessment and research*. New York, NY: Dryden.
- Edwards, W. (1954). The theory of decision making. *Psychological bulletin*, 51, 380-417. <http://dx.doi.org/10.1037/h0053870>
- Edwards, A. L. (1959). *Edwards personal preference schedule (2nd ed.)*. Oxford, England: Psychological Corp.
- Ellingson, J. E. (2012). People fake only when they need to fake. En M. Ziegler, C. MacCann y R. D. Roberts (Eds.), *New perspectives on faking in personality assessment* (pp. 34-52). New York, NY, US: Oxford University Press.
- Ellingson, J. E. y McFarland, L. A. (2011). Understanding faking behavior through the lens of motivation: An application of VIE theory. *Human Performance*, 24, 322-337. <https://doi.org/10.1080/08959285.2011.597477>
- Ellingson, J. E., Sackett, P. R. y Connelly, B. S. (2007). Personality assessment across selection and development contexts: Insights into response distortion. *Journal of Applied Psychology*, 92, 386-395. <http://dx.doi.org/10.1037/0021-9010.92.2.386>
- Ellingson, J. E., Sackett, P. R. y Hough, L. M. (1999). Social desirability corrections in personality measurement: Issues of applicant comparison and construct validity. *Journal of Applied Psychology*, 84, 155-166. <http://dx.doi.org/10.1037/0021-9010.84.2.155>

- Ellingson, J. E., Smith, D. B. y Sackett, P. R. (2001). Investigating the influence of social desirability on personality factor structure. *Journal of Applied Psychology*, 86, 122-133.
- English, A., Griffith, R. L., Graseck, M. y Steelam, L. A. (2005). Frame of reference, applicant faking and the predictive validity of non-cognitive measures: A matter of context. Manuscrito pendiente de publicación.
- Eysenck, H. J. (1959). *Manual of the Maudsley personality inventory*. University of London Press.
- Eysenck, H. J. y Eysenck, S. B. G. (1975). *Manual of the Eysenck Personality Questionnaire (junior and adult)*. Hodder and Stoughton.
- Fandt, P. M. y Ferris, G. R. (1990). The management of information and impressions: When employees behave opportunistically. *Organizational Behavior and Human Decision Processes*, 45, 140-158. [https://doi.org/10.1016/0749-5978\(90\)90008-W](https://doi.org/10.1016/0749-5978(90)90008-W)
- Fekken, G. C. y Holden, R. R. (1992). Response latency evidence for viewing personality traits as schema indicators. *Journal of Research in Personality*, 26, 103-120. [https://doi.org/10.1016/0092-6566\(92\)90047-8](https://doi.org/10.1016/0092-6566(92)90047-8)
- Fine, S. y Pirak, M. (2016). Faking fast and slow: Within-person response time latencies for measuring faking in personnel testing. *Journal of Business and Psychology*, 31, 51-64. <https://doi.org/10.1007/s10869-015-9398-5>
- \*Fineman, S. (1975). The Work Preference Questionnaire: A measure of managerial need for achievement. *Journal of Occupational Psychology*, 48, 11-32. <https://doi.org/10.1111/j.2044-8325.1975.tb00293.x>
- Fisher, P. A., Robie, C., Christiansen, N. D., Speer, A. B. y Schneider, L. (2019). Criterion-related Validity of Forced-Choice Personality Measures: A Cautionary Note Regarding Thurstonian IRT versus Classical Test Theory Scoring. *Personnel Assessment and Decisions*, 5, 49-61. <https://doi.org/10.25035/pad.2019.01.003>
- Fiske, D. W. (1949). Consistency of the factorial structures of personality ratings from different sources. *The Journal of Abnormal and Social Psychology*, 44, 329-344. <http://dx.doi.org/10.1037/h0057198>
- \*Ford, K. A. y Gunderson, E. K. (1962). *Personality characteristics (EPPS) of Antarctic volunteers* (Nº NMNRU-62-18). Navy Medical Neuropsychiatric Research Unit, San Diego, CA.

- Furnham, A. (1986). Response bias, social desirability and dissimulation. *Personality and Individual Differences*, 7, 385-400. [https://doi.org/10.1016/0191-8869\(86\)90014-0](https://doi.org/10.1016/0191-8869(86)90014-0)
- Furnham, A. (1990a). *The Protestant work ethic: The psychology of work-related beliefs and behaviors*. New York, NY: Routledge.
- Furnham, A. (1990b). Faking personality questionnaires: Fabricating different profiles for different purposes. *Current Psychology*, 9, 46-55. <https://doi.org/10.1007/BF02686767>
- \*Furnham, A. y Craig, S. (1987). Fakeability and correlates of the perception and preference inventory. *Personality and Individual Differences*, 8, 459-470. [https://doi.org/10.1016/0191-8869\(87\)90207-8](https://doi.org/10.1016/0191-8869(87)90207-8)
- Garry, R. (1953). Individual differences in ability to fake vocational interests. *Journal of Applied Psychology*, 37, 33-37. <http://dx.doi.org/10.1037/h0058887>
- Goffin, R. D. y Boyd, A. C. (2009). Faking and personality assessment in personnel selection: Advancing models of faking. *Canadian Psychology / Psychologie Canadienne*, 50, 151-160. <http://dx.doi.org/10.1037/a0015946>
- Goffin, R. D. y Christiansen, N. D. (2003). Correcting personality tests for faking: A review of popular personality tests and an initial survey of researchers. *International Journal of Selection and Assessment*, 11, 340-344. <https://doi.org/10.1111/j.0965-075X.2003.00256.x>
- Goldberg, L. R. (1981). Language and individual differences: The search for universals in personality lexicons. En L. Wheeler (Ed), *Review of Personality and Social Psychology* (págs. 141-165). Beverly Hills, CA: Sage.
- Goldberg, L. R. (1992). The development of markers for the Big-Five factor structure. *Psychological Assessment*, 4, 26-42.
- Gordon, L. V. (1951). Validities of the forced-choice and questionnaire methods of personality measurement. *Journal of Applied Psychology*, 35, 407-412. <http://dx.doi.org/10.1037/h0058853>
- Gordon, M. J. (1963). Optimal investment and financing policy. *The Journal of Finance*, 18, 264-272. <http://dx.doi.org/10.2307/2977907>
- Gordon, L. V. (1993). *Manual: Gordon Personal Profile-Inventory*. San Antonio: The Psychological Corporation.
- Gordon, L. V. y Chevrier, J. M. (1993). *Gordon personal profile-inventory*. Psychological Corporation.

- \*Gordon, L. V. y Stapleton, E. S. (1956). Fakability of a forced-choice personality test under realistic high school employment conditions. *Journal of Applied Psychology*, 40, 258-262. <http://dx.doi.org/10.1037/h0043595>
- Gough, H. G. (1957) *The California Psychological Inventory administrator's guide*. Palo Alto, CA: CPP.
- Gough, H. G. (1975). Personality factors related to reported severity of menstrual distress. *Journal of Abnormal Psychology*, 84, 59-65. <http://dx.doi.org/10.1037/h0076260>
- Graham, M. A., Monday, J., O'Brien, K. y Steffen, S. (1994). Cheating at small colleges: an examination of student and faculty attitudes and behaviors. *Journal of College Student Development*, 35, 255-260.
- Granello, D. H. y Wheaton, J. E. (1998). Self-perceived multicultural competencies of African American and European American vocational rehabilitation counselors. *Rehabilitation Counseling Bulletin*, 42, 2-15.
- Grieve, R. y McSwiggan, C. (2014). La predicción de la intención de falsear las respuestas en los tests psicológicos: ¿qué creencias normativas son importantes? *Journal of Work and Organizational Psychology*, 30, 23-28. <http://dx.doi.org/10.5093/tr2014a3>
- Griffith, R. L., Chmielowski, T. y Yoshita, Y. (2007). Do applicants fake? An examination of the frequency of applicant faking behavior. *Personnel Review*, 36, 341-355.
- Griffith, R. L. y Converse, P. D. (2012). The rules of evidence and the prevalence of applicant faking. En M. Ziegler, C. MacCann y R. D. Roberts (Eds.), *New Perspectives on Faking in Personality Assessment* (págs. 34-52). New York, NY: Oxford University Press.
- Griffith, R. L., Lee, L. M., Peterson, M. H. y Zickar, M. J. (2011). First dates and little white lies: A trait contract classification theory of applicant faking behavior. *Human Performance*, 24, 338-357. <https://doi.org/10.1080/08959285.2011.597475>
- Griffith, R. L., Malm, T., English, A., Yoshita, Y. y Gujar, A. (2006). Applicant faking behavior: Teasing apart the influence of situational variance, cognitive biases, and individual differences. En R. L. Griffith y M. H. Peterson (Eds.), *A Closer Examination of Applicant Faking Behavior* (págs. 151-178). Greenwich, CT: Information Age.
- Griffith, R. L. y McDaniel, M. (2006). The nature of deception and applicant faking behavior. En R. L. Griffith y M. H. Peterson (Eds.), *A Closer Examination of Applicant Faking Behavior* (págs. 113-148). Greenwich, CT: Information Age.

- Griffith, R. L. y Peterson, M. H. (2006). *A closer examination of applicant faking behavior*. Greenwich, CT: Information Age Publishing, Inc.
- Griffith, R. L. y Peterson, M. H. (2011). One piece at a time: The puzzle of applicant faking and a call for theory. *Human Performance*, 24, 291-301. <https://doi.org/10.1080/08959285.2011.597474>
- Griffith, R. L., Peterson, M. H., Quist, J., Benda, A. y Evans, A. L. (2008, abril). Faking the personality profile: Easier said than done. En R. L. Griffith y M. H. Peterson (Chairs), *Complex problems, simple solutions: Contemporary research in applicant faking behavior*. Symposium presentado en la 23ª Annual Conference for the Society for Industrial and Organizational Psychology, San Francisco, CA.
- Griffith, R. L. y Robie, C. (2013) Personality testing and the “F-Word”: Revisiting seven questions about faking. En N. D. Christiansen y R. P. Tett (Eds.), *Handbook of Personality at Work* (págs. 253-280). New York, NY: Taylor & Francis.
- Guion, R. M. y Gottier, R. F. (1965). Validity of personality measures in personnel selection. *Personnel Psychology*, 18, 135-164. <https://doi.org/10.1111/j.1744-6570.1965.tb00273.x>
- Hall, R. C. W. y Hall, R. C. W. (2011). Plaintiffs who malingering: Impact on litigation on fake testimony. En M. Ziegler, C. MacCann y R. Roberts (Eds.), *New Perspectives on Faking in Personality Assessment* (págs. 255-281). New York, NY: Oxford University Press.
- Hathaway, S. R. y McKinley, J. C. (1951). *Manual for the Minnesota Multiphasic Personality Inventory (Rev. ed.)*. New York, NY: Psychological Corporation.
- \*Hedberg, R. (1962). More on forced-choice test fakability. *Journal of Applied Psychology*, 46, 125-127. <http://dx.doi.org/10.1037/h0038453>
- Heggestad, E. D. (2012). A conceptual representation of faking: Putting the horse back in front of the cart. En M. Ziegler, C. MacCann y R. Roberts (Eds.), *New Perspectives on Faking in Personality Assessment* (págs. 87-101). New York, NY: Oxford University Press
- Heggestad, E. D., George, E. y Reeve, C. L. (2006). Transient error in personality scores: Considering honest and faked responses. *Personality and Individual Differences*, 40, 1201-1211. <https://doi.org/10.1016/j.paid.2005.10.014>
- \*Heggestad, E. D., Morrison, M., Reeve, C. L. y McCloy, R. A. (2006). Forced-choice assessments of personality for selection: Evaluating issues of normative assessment and faking resistance. *Journal of Applied Psychology*, 91, 9. <http://dx.doi.org/10.1037/0021-9010.91.1.9>



- Heller, M. (2005). Court ruling that employer's integrity test violated ADA could open door to litigation. *Workforce Management*, 84, 74-77.
- Helmes, E. y Holden, R. R. (2003). The construct of social desirability: one or two dimensions? *Personality and Individual Differences*, 34, 1015-1023. [https://doi.org/10.1016/S0191-8869\(02\)00086-7](https://doi.org/10.1016/S0191-8869(02)00086-7)
- \*Helton, K. T. y Street Jr, D. R. (1992). *The five-factor personality model and naval aviation candidates* (Nº NAMRL-1379). Naval Aerospace Medical Research Lab, Pensacola, FL.
- Hendriks, A. A. J., Hofstee, W. K. B. y De Raad, B. (1999). The Five-Factor Personality Inventory (FFPI). *Personality and Individual Differences*, 27, 307-325. doi:10.1016/S0191-8869(98)00245-1
- Hicks, L. E. (1970). Some properties of ipsative, normative, and forced-choice normative measures. *Psychological Bulletin*, 74, 167-184. <http://dx.doi.org/10.1037/h0029780>
- \*Hirsh, J. B. y Peterson, J. B. (2008). Predicting creativity and academic success with a "fake-proof" measure of the Big Five. *Journal of Research in Personality*, 42, 1323-1333. <https://doi.org/10.1016/j.jrp.2008.04.006>
- Hogan, R. (1983). A socioanalytic theory of personality. En M. Page y R. Dienstbier (Eds.), *Nebraska Symposium on Motivation, 1982* (págs. 55-89). Lincoln: University of Nebraska Press.
- Hogan, R. (1986). *Manual for the Hogan Personality Inventory*. Minneapolis, MN: National Computer Systems.
- Hogan, R. T. (1991). Personality and personality measurement. En M. D. Dunnette y L. M. Hough (Eds.), *Handbook of Industrial and Organizational Psychology* (págs. 873-919). Palo Alto, CA: Consulting Psychologists Press..
- Hogan, J., Barrett, P. y Hogan, R. (2007). Personality measurement, faking, and employment selection. *Journal of Applied Psychology*, 92, 1270-1285. <http://dx.doi.org/10.1037/0021-9010.92.5.1270>
- Holden, R. R. y Book, A. S. (2012). Faking does distort self-report personality assessment. En M. Ziegler, C. MacCann y R. D. Roberts (Eds.), *New Perspectives on Faking in Personality Assessment* (págs. 71-84). New York, NY: Oxford University Press.
- Holden, R. R., Kroner, D. G., Fekken, G. C. y Popham, S. M. (1992). A model of personality test item response dissimulation. *Journal of Personality and Social Psychology*, 63, 272-279. <http://dx.doi.org/10.1037/0022-3514.63.2.272>

- Hooper, A. C. (2007). *Self-presentation on personality measures in lab and field settings: A meta-analysis* (Tesis doctoral no publicada). University of Minnesota, Minneapolis, MN.
- Horn, J. L. (1971). Motivation and dynamic calculus concepts from multivariate experiment. En R. B. Cattell (Ed.) *Handbook of Multivariate Experimental Psychology* (2ª Ed.) (págs. 611-641). Chicago, IL: Tand McNally,
- Hough, L. M. (1998a). Effects of intentional distortion in personality measurement and evaluation of suggested palliatives. *Human Performance*, 11, 209-244. <https://doi.org/10.1080/08959285.1998.9668032>
- Hough, L. (1998b). Personality at work: issues and evidence. En M. D. Hakel (Ed.), *Beyond multiple choice: evaluating alternatives to traditional testing for selection* (págs. 131-166). Mahwah, NJ: Lawrence Erlbaum Associates.
- Hough, L. M., Eaton, N. K., Dunnette, M. D., Kamp, J. D. y McCloy, R. A. (1990). Criterion-related validities of personality constructs and the effect of response distortion on those validities. *Journal of Applied Psychology*, 75, 581. <http://dx.doi.org/10.1037/0021-9010.75.5.581>
- Hough, L. M., Kamp, J. y Ashworth, S. D. (1993). *Development of "Assessment of Background and Life Experiences" (ABLE): A temperament inventory*. Minneapolis, MN: Personnel Decisions Research Institutes, Inc.
- Hough, L. M. y Ones, D. S. (2002). The structure, measurement, validity, and use of personality variables in industrial, work, and organizational psychology. En N. Anderson, D. S. Ones, H. K. Sinangil y C. Viswesvaran (Eds.), *Handbook of industrial, work and organizational psychology (Vol. 1)* (págs. 233-277). Thousand Oaks, CA: Sage Publications Ltd.
- Hsu, L. M., Santelli, J. y Hsu, J.R. (1989). Faking detection validity and incremental validity of response latencies to MMPI subtle and obvious items. *Journal of Personality Assessment*, 53, 278-295. [https://doi.org/10.1207/s15327752jpa5302\\_6](https://doi.org/10.1207/s15327752jpa5302_6)
- Hu, L. T. y Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6, 1-55. <https://doi.org/10.1080/10705519909540118>
- Humm, D. G. y Wadsworth, G. W. (1934). *The Humm-Wadsworth temperament scale*. Los Angeles, CA: Humm and Associates.
- Hunt, E. B. (2011). *Human intelligence*. Cambridge: Cambridge University Press.



- Hunter, J. E. y Schmidt, F. L. (1987a). *Error in the meta-analysis of correlations: The mean correlation*. Manuscrito no publicado, Departamento de Psicología, Michigan State University, Michigan, Estados Unidos.
- Hunter, J. E. y Schmidt, F. L. (1987b). *Error in the meta-analysis of correlations: The standard deviation*. Manuscrito no publicado, Departamento de Psicología, Michigan State University, Michigan, Estados Unidos.
- Hunter, J. E. y Schmidt, F. L. (2004). *Methods of Meta-Analysis: Correcting Error and Bias in Research Findings* (2nd ed.). Thousand Oaks, CA: Sage.
- Hurtz, G. M. y Donovan, J. J. (2000). Personality and job performance: The Big Five revisited. *Journal of Applied Psychology*, 85, 869-879.
- Ingold, P. V., Kleinmann, M., König, C. J. y Melchers, K. G. (2015). Shall we continue or stop disapproving of self-presentation? Evidence on impression management and faking in a selection context and their relation to job performance. *European Journal of Work and Organizational Psychology*, 24, 420-432. <https://doi.org/10.1080/1359432X.2014.915215>
- Inwald, R., Knatz, H. y Shusman, L. (1983). *Inwald Personality Inventory manual*. New York: Hilson Research Jackson, D.N. (1999) *Personality Research Form Manual* (3rd ed.). Port Huron, MI: Sigma Assessment Systems.
- \*Jackson, D. N., Wroblewski, V. R. y Ashton, M. C. (2000). The impact of faking on employment tests: Does forced choice offer a solution? *Human Performance*, 13, 371-388. [https://doi.org/10.1207/S15327043HUP1304\\_3](https://doi.org/10.1207/S15327043HUP1304_3)
- James, L. R. (1998). Measurement of personality via conditional reasoning. *Organizational Research Methods*, 1, 131-163. <https://doi.org/10.1177/109442819812001>
- Jensen, A. R. (1985). Methodological and statistical techniques for the chronometric study of mental abilities. En C. R. Reynolds y V. L. Wilson (Eds.), *Methodological and statistical advances in the study of individual differences*. New York, NY: Plenum.
- Jeong, Y. R., Christiansen, N. D., Robie, C., Kung, M. C. y Kinney, T. B. (2017). Comparing applicants and incumbents: Effects of response distortion on mean scores and validity of personality measures. *International Journal of Selection and Assessment*, 25, 311-315. <https://doi.org/10.1111/ijsa.12182>
- Jiang, G., Mai, Y. y Yuan, K. H. (2017). Advances in measurement invariance and mean comparison of latent variables: equivalence testing and a projection-based approach. *Frontiers in Psychology*, 8, 1823. <https://doi.org/10.3389/fpsyg.2017.01823>

- Johnson, J. A. y Hogan, R. (2006). A socioanalytic view of faking. En R. L. Griffith y M. H. Peterson (Eds.), *A closer examination of applicant faking behavior* (págs. 209-231). Greenwich, CT: Information Age.
- Johnson, C. E., Wood, R. y Blinkhorn, S. F. (1988). Spuriouser and spuriouser: The use of ipsative personality tests. *Journal of Occupational Psychology*, 61, 153-162. <https://doi.org/10.1111/j.2044-8325.1988.tb00279.x>
- Jonason, P. K., Lyons, M., Baughman, H. M. y Vernon, P. A. (2014). What a tangled web we weave: The Dark Triad traits and deception. *Personality and Individual Differences*, 70, 117-119. <https://doi.org/10.1016/j.paid.2014.06.038>
- Jöreskog, K. y Sörbom, D. (1998). LISREL 8.20 and PRELIS 2.20 [programa informático]. Chicago: Scientific Software Inc.
- Judge, T. A. y Bono, J. E. (2001). Relationship of core self-evaluations traits-self-esteem, generalized self-efficacy, locus of control, and emotional stability with job satisfaction and job performance: A meta-analysis. *Journal of Applied Psychology*, 86, 80-92. <https://doi.org/10.1037/0021-9010.86.1.80>
- Kacmar, K. M., Carlson, D. S. y Bratton, V. K. (2004). Situational and dispositional factors as antecedents of ingratiation behaviors in organizational settings. *Journal of Vocational Behavior*, 65, 309-331. <https://doi.org/10.1016/j.jvb.2003.09.002>
- \*Kaess, W. A. y Witryol, S. L. (1957). Positive and negative faking on a forced-choice authoritarian scale. *Journal of Applied Psychology*, 41, 333-339. <http://dx.doi.org/10.1037/h0043451>
- Kahneman, D., Slovic, S. P., Slovic, P. y Tversky, A. (1982). *Judgment under uncertainty: Heuristics and biases*. Cambridge university press.
- Kanfer, R. y Ackerman, P. L. (1989). Motivation and cognitive abilities: An integrative/aptitude-treatment interaction approach to skill acquisition. *Journal of Applied Psychology*, 74, 657-690. <http://dx.doi.org/10.1037/0021-9010.74.4.657>
- Kelly, E. L., Miles, C. C. y Terman, L. M. (1936). Ability to influence one's score on a typical pencil-and-paper test of personality. *Character & Personality; A Quarterly for Psychodiagnostic & Allied Studies*. <http://dx.doi.org/10.1111/j.1467-6494.1936.tb02123.x>
- Kim, B. H. (2011). Deception and applicant faking: Putting the pieces together. En G. P. Hodgkinson y J. K. Ford (Eds.), *International Review of Industrial and Organizational Psychology* (págs. 239-292). New York, NY: John Wiley & Sons, Ltd.

- \*Kirchner, W. K. (1962). " Real-life" faking on the Edwards Personal Preference Schedule by sales applicants. *Journal of Applied Psychology*, 46, 128-130. <http://dx.doi.org/10.1037/h0039528>
- \*Kirchner, W. K., Dunnette, M. D. y Mousley, N. (1960). Use of the Edwards Personal Preference Schedule in the selection of salesmen. *Personnel Psychology*, 421-424. <http://dx.doi.org/10.1111/j.1744-6570.1960.tb02099.x>
- Kluger, A. N. y Colella, A. (1993). Beyond the mean bias: The effect of warning against faking on biodata item variances. *Personnel Psychology*, 46, 763-780. <https://doi.org/10.1111/j.1744-6570.1993.tb01568.x>
- Komar, S., Brown, D. J., Komar, J. A. y Robie, C. (2008). Faking and the validity of conscientiousness: A Monte Carlo investigation. *Journal of Applied Psychology*, 93, 140–154. <https://doi.org/10.1037/0021-9010.93.1.140>
- Komar, S., Komar, J. A., Robie, C. y Taggar, S. (2010). Speeding personality measures to reduce faking: A self-regulatory model. *Journal of Personnel Psychology*, 9, 126-137. <https://doi.org/10.1027/1866-5888/a000016>
- Krammer, R. (2017). *Using personality measures for selection decisions: predictive utility and applicants' faking behavior*. Tesis doctoral.
- \*Krug, R. E. (1958). The effect of specific selection sets on a forced-choice self-description inventory. *Journal of Applied Psychology*, 42, 89-92. <http://dx.doi.org/10.1037/h0044420>
- Kuncel, N. R. y Borneman, M. J. (2007). Toward a new method of detecting deliberately faked personality tests: The use of idiosyncratic item responses. *International Journal of Selection and Assessment*, 15, 220-231. <https://doi.org/10.1111/j.1468-2389.2007.00383.x>
- Kuncel, N. R., Goldberg, L. R. y Kiger, T. (2011). A plea for process in personality prevarication. *Human Performance*, 24, 373-378. <https://doi.org/10.1080/08959285.2011.597476>
- Kuncel, N. R., Hezlett, S. A. y Ones, D. S. (2004). Academic performance, career potential, creativity, and job performance: Can one construct predict them all? *Journal of Personality and Social Psychology*, 86, 148-161. <http://dx.doi.org/10.1037/0022-3514.86.1.148>
- LaHuis, D. M. (2005). Individual differences in applicant reactions: A job-search perspective. *International Journal of Selection and Assessment*, 13, 150-159. <https://doi.org/10.1111/j.0965-075X.2005.00309.x>

- Landauer, T. K. y Dumais, S. T. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review*, 104, 211-240.
- Landers, R. N., Sackett, P. R. y Tuzinski, K. A. (2011). Retesting after initial failure, coaching rumors, and warnings against faking in online personality measures for selection. *Journal of Applied Psychology*, 96, 202-210.
- \* Larson, N. L., Lewis, R. J., O'Neill, T. A. y Carswell, J. J. (2013, April). *Are forced choice personality measures contaminated by general mental ability?* Poster presentado en la 28<sup>th</sup> Annual Conference of the Society for Industrial and Organizational Psychology, Houston, TX.
- Leary, M. R. y Kowalski, R. M. (1990). Impression management: A literature review and two-component model. *Psychological Bulletin*, 107, 34-47.
- LeBreton, J. M., Barksdale, C. D., Robin, J. y James, L. R. (2007). Measurement issues associated with conditional reasoning tests: Indirect measurement and test faking. *Journal of Applied Psychology*, 92, 1-16. <http://dx.doi.org/10.1037/0021-9010.92.1.1>
- Lee, P., Lee, S. y Stark, S. (2018). Examining validity evidence for multidimensional forced choice measures with different scoring approaches. *Personality and Individual Differences*, 123, 229-235. <https://doi.org/10.1016/j.paid.2017.11.031>
- Levashina, J. y Campion, M. A. (2007). Measuring faking in the employment interview: development and validation of an interview faking behavior scale. *Journal of Applied Psychology*, 92, 1638-1656. <http://dx.doi.org/10.1037/0021-9010.92.6.1638>
- Levashina, J., Morgeson, F. P. y Campion, M. A. (2009). They don't do it often, but they do it well: Exploring the relationship between applicant mental abilities and faking. *International Journal of Selection and Assessment*, 17, 271-281. <https://doi.org/10.1111/j.1468-2389.2009.00469.x>
- Levin, R. A. y Zickar, M. J. (2002). Investigating self-presentation, lies, and bullshit: Understanding faking and its effects on selection decisions using theory, field research, and simulation. En J. M. Brett y F. Drasgow (Eds.), *The Psychology of Work. Theoretically Based Empirical Research* (págs. 253-276). Mahwah: Lawrence Erlbaum Associates, Inc.
- Li, A. y Bagger, J. (2006). Using the BIDR to distinguish the effects of impression management and self-deception on the criterion validity of personality measures: A meta-analysis. *International Journal of Selection and Assessment*, 14, 131-141. <https://doi.org/10.1111/j.1468-2389.2006.00339.x>

- Lorenzo-Seva, U. y Ferrando, P. J. (2018). FACTOR (10.8. 02) [Software]. Tarragona, España: Universidad Rovira i Virgili.
- Lueke, S. B., Snell, A. F., Illingworth, A. J. y Paidas, S. M. (2001). *An empirical test of an interactional model of faking*. Comunicación presentada en el Meeting of the Society for Industrial and Organizational Psychology, San Diego, CA.
- \*Mahar, D., Coburn, B., Griffin, N., Hemeter, F., Potappel, C., Turton, M. y Mulgrew, K. (2006). Stereotyping as a response strategy when faking personality questionnaires. *Personality and Individual Differences*, 40, 1375-1386. <https://doi.org/10.1016/j.paid.2005.11.018>
- \*Mahar, D., Cologon, J. y Duck, J. (1995). Response strategies when faking personality questionnaires in a vocational selection setting. *Personality and Individual Differences*, 18, 605-609. [https://doi.org/10.1016/0191-8869\(94\)00200-C](https://doi.org/10.1016/0191-8869(94)00200-C)
- Marcus, B. (2006). Relationships between faking, validity, and decision criteria in personnel selection. *Psychology Science*, 48, 226-246.
- Marcus, B. (2009). 'Faking' from the applicant's perspective: A theory of self-presentation in personnel selection settings. *International Journal of Selection and Assessment*, 17, 417-430. <https://doi.org/10.1111/j.1468-2389.2009.00483.x>
- Marshall, M. B., De Fruyt, F., Rolland, J. P. y Bagby, R. M. (2005). Socially desirable responding and the factorial stability of the NEO PI-R. *Psychological Assessment*, 17, 379-384. <http://dx.doi.org/10.1037/1040-3590.17.3.379>
- McCloy, R. A., Heggstad, E. D. y Reeve, C. L. (2005). A silk purse from the sow's ear: Retrieving normative information from multidimensional forced-choice items. *Organizational Research Methods*, 8, 222-248. <https://doi.org/10.1177/1094428105275374>
- McCrae, R. R. y Costa, P. T. (1983). Social desirability scales: More substance than style. *Journal of Consulting and Clinical Psychology*, 51, 882-888. <http://dx.doi.org/10.1037/0022-006X.51.6.882>
- McCrae, R. R. y Costa, P. T. (1985). Updating Norman's" adequacy taxonomy": Intelligence and personality dimensions in natural language and in questionnaires. *Journal of Personality and Social Psychology*, 49, 710-721.
- McCrae, R. R. y Costa, P. T. (1987). Validation of the Five-Factor Model of personality across instruments and observers. *Journal of Personality and Social Psychology*, 52, 81-90. <http://dx.doi.org/10.1037/0022-3514.52.1.81>
- McDaniel, M. A., Douglas, E. F. y Snell, A. F. (1997, abril). *A survey of deception among job seekers*. Comunicación presentada en la 12ª Annual Conference of the Society for Industrial and Organizational Psychology, St. Louis, MO.



- McDaniel, M. A. y Timm, H. (1990, agosto). *Lying takes time: Predicting deception in biodata using response latency*. Comunicación presentada en la 98ª Annual Convention of the American Psychological Association, Boston.
- McDaniel, M. A., Whetzel, D. L., Schmidt, F. L. y Maurer, S. D. (1994). The validity of employment interviews: A comprehensive review and meta-analysis. *Journal of Applied Psychology*, 79, 599-616. <http://dx.doi.org/10.1037/0021-9010.79.4.599>
- McDougall, W. (1932). Of the words character and personality. *Journal of Personality*, 1, 3-16.
- McFarland, L. A. (2003). Warning against faking on a personality test: Effects on applicant reactions and personality tests scores. *International Journal of Selection and Assessment*, 11, 265-276. <https://doi.org/10.1111/j.0965-075X.2003.00250.x>
- McFarland, L. A. y Ryan, A. M. (2000). Variance in faking across noncognitive measures. *Journal of Applied Psychology*, 85, 812-821. <http://dx.doi.org/10.1037/0021-9010.85.5.812>
- McFarland, L. A. y Ryan, A. M. (2006). Toward an integrated model of applicant faking behavior. *Journal of Applied Social Psychology*, 36, 979-1016. <https://doi.org/10.1111/j.0021-9029.2006.00052.x>
- McLarnon, M. J., DeLongchamp, A. C. y Schneider, T. J. (2019). Faking it! Individual differences in types and degrees of faking behavior. *Personality and Individual Differences*, 138, 88-95. <https://doi.org/10.1016/j.paid.2018.09.024>
- Meade, A.W. (2004) Psychometric problems and issues involved with creating and using ipsative measures for selection. *Journal of Occupational and Organizational Psychology*, 77, 531-552.
- Mealey, L. (1995). The sociobiology of sociopathy: An integrated evolutionary model. *Behavioral and Brain sciences*, 18, 523-541. DOI: <https://doi.org/10.1017/S0140525X00039595>
- Meehl, P. E. y Hathaway, S. R. (1946). The K factor as a suppressor variable in the Minnesota Multiphasic Personality Inventory. *Journal of Applied Psychology*, 30, 525-564. <https://psycnet.apa.org/doi/10.1037/h0053634>
- Mersman, J. L. y Shultz, K. S. (1998). Individual differences in the ability to fake on personality measures. *Personality and Individual Differences*, 24, 217-227. [https://doi.org/10.1016/S0191-8869\(97\)00160-8](https://doi.org/10.1016/S0191-8869(97)00160-8)
- Mesmer-Magnus, J. y Viswesvaran, C. (2006). Assessing response distortion in personality tests. En R. L. Griffith y M. H. Peterson (Eds.), *A closer examination of applicant faking behavior* (págs. 85-113). Greenwich, CT: Information Age Publishing, Inc.

- Mesmer-Magnus, J., Viswesvaran, C., Deshpande, S. y Joseph, J. (2006). Social desirability: The role of over-claiming, self-esteem, and emotional intelligence. *Psychology Science*, 48, 336-356.
- Messick, S. (1960). Dimensions of social desirability. *Journal of Consulting Psychology*, 24, 279-287. <http://dx.doi.org/10.1037/h0044153>
- Michaelis, W. y Eysenck, H. J. (1971). The determination of personality inventory factor patterns and intercorrelations by changes in real-life motivation. *The Journal of Genetic Psychology*, 118, 223-234. <https://doi.org/10.1080/00221325.1971.10532611>
- Miller, C. E. y Barrett, G. V. (2008). The coachability and fakability of personality-based selection tests used for police selection. *Public Personnel Management*, 37, 339-351. <https://doi.org/10.1177/009102600803700306>
- Millsap, R. E. (2011). *Statistical Approaches to Measurement Invariance*. New York, NY: Routledge.
- Morgeson, F. P., Campion, A. M., Dipboye R. L., Hollenbeck R. J., Murphy K. y Schmitt, N. (2007). Are we getting fooled again? Coming to terms with limitations in the use of personality tests for personnel selection, *Personnel Psychology*, 60, 1029-1049. <https://doi.org/10.1111/j.1744-6570.2007.00100.x>
- \*Mourer, S. A. (1959). *A statistical analysis of the relations between interpersonal perception and adjustment in leaders*. Tesis de máster no publicada. University of Nebraska at Omaha, Nebraska, NE.
- Mueller-Hanson, R., Heggstad, E. D. y Thornton, G. C. III. (2003). Faking and selection: Considering the use of personality from select-in and select-out perspectives. *Journal of Applied Psychology*, 88, 348-355. <http://dx.doi.org/10.1037/0021-9010.88.2.348>
- Mueller-Hanson, R. A., Heggstad, E. D. y Thornton, G. C. (2006). Individual differences in impression management: An exploration of the psychological processes underlying faking. *Psychology Science*, 48, 288-312.
- Myers, I. B., McCaulley, M. H., Quenk, N. L. y Hammer, A. L. (1998). *MBTI manual: A guide to the development and use of the Myers-Briggs Type Indicator* (Vol. 3). Palo Alto, CA: Consulting Psychologists Press.
- Nederhof, A. J. (1985). Methods of coping with social desirability bias: a review. *European journal of social psychology*, 15, 263-280.
- Newman, M. L., Pennebaker, J. W., Berry, D. S. y Richards, J. M. (2003). Lying words: Predicting deception from linguistic styles. *Personality and Social Psychology Bulletin*, 29, 665-675. <https://doi.org/10.1177/0146167203029005010>



- Newstead, S. E., Franklyn-Stokes, A. y Armstead, P. (1996). Individual differences in student cheating. *Journal of Educational Psychology*, 88, 229-241.
- Norman, W. T. (1963a). Toward an adequate taxonomy of personality attributes: Replicated factor structure in peer nomination personality ratings. *The Journal of Abnormal and Social Psychology*, 66, 574-583. <http://dx.doi.org/10.1037/h0040291>
- \*Norman, W. T. (1963b). Personality measurement, faking, and detection: An assessment method for use in personnel selection. *Journal of Applied Psychology*, 47, 225-241. <http://dx.doi.org/10.1037/h0042106>
- Nguyen, N. T. y McDaniel, M. A. (2000, diciembre). *Brain size and intelligence: A meta-analysis*. Comunicación presentada en la 1ª Annual Conference of the International Society of Intelligence Research, Cleveland, OH.
- O'Connell, M. S., Kung, M. C. y Tristan, E. (2011). Beyond impression management: Evaluating three measures of response distortion and their relationship to job performance. *International Journal of Selection and Assessment*, 19, 340-351. <https://doi.org/10.1111/j.1468-2389.2011.00563.x>
- \*O'Neill, T. A., Lewis, R. J., Law, S. J., Larson, N., Hancock, S., Radan, J., Lee, N. y Carswell, J. J. (2017). Forced-choice pre-employment personality assessment: Construct validity and resistance to faking. *Personality and Individual Differences*, 115, 120-127. <https://doi.org/10.1016/j.paid.2016.03.075>
- Ones, D. S. y Viswesvaran, C. (1998). The effects of social desirability and faking on personality and integrity assessment for personnel selection. *Human Performance*, 11, 245-269. <https://doi.org/10.1080/08959285.1998.9668033>
- Ones, D. S. y Viswesvaran, C. (1999). Relative importance of personality dimensions for expatriate selection: A policy capturing study. *Human Performance*, 12, 275-294. <https://doi.org/10.1080/08959289909539872>
- Ones, D. S., Viswesvaran, C. y Reiss, A. D. (1996). Role of social desirability in personality testing for personnel selection: The red herring. *Journal of Applied Psychology*, 81, 660-679.
- Ones, D. S., Viswesvaran, C. y Schmidt, F. L. (1993). Comprehensive meta-analysis of integrity test validities: Findings and implications for personnel selection and theories of job performance. *Journal of Applied Psychology*, 78, 679-703. <https://doi.org/10.1037/0021-9010.78.4.679>
- Pannone, R. D. (1984). Predicting test performance: A content valid approach to screening applicants. *Personnel Psychology*, 37, 507-514. <https://doi.org/10.1111/j.1744-6570.1984.tb00526.x>

- Paulhus, D. L. (1981). Control of social desirability in personality inventories: Principal-factor deletion. *Journal of Research in Personality*, 15, 383-388. [https://doi.org/10.1016/0092-6566\(81\)90035-0](https://doi.org/10.1016/0092-6566(81)90035-0)
- Paulhus, D. L. (1984). Two-component models of socially desirable responding. *Journal of Personality and Social Psychology*, 46, 598-609. <http://dx.doi.org/10.1037/0022-3514.46.3.598>
- Paulhus, D. L. (1986). Self-deception and impression management in test responses. En A. Angleitner y J. S. Wiggins (Eds.) *Personality assessment via questionnaires* (págs. 143-165). Berlin, Heidelberg: Springer.
- Paulhus, D. L. (1988). *Assessing self-deception and impression management in self-reports: The Balanced Inventory of Desirable Responding*. University of British Columbia, Vancouver, British Columbia, Canada.
- Paulhus, D. L. (1991). *Measurement and control of response bias*. En J. Robinson, P. R. Shaver y L. S. Wrightsman (Eds.), *Measures of Personality and Social Psychological Attitudes* (págs. 17-59). New York, NY: Academic Press.
- Paulhus, D. L. (2002). Socially desirable responding: The evolution of a construct. En H. Braun, D. N. Jackson y D. E. Wiley (Eds.), *The Role of Constructs in Psychological and Educational Measurement* (págs. 67-88). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Paulhus, D. L. (2012). Overclaiming on Personality Questionnaire. En M. Ziegler, C. MacCan y R. D. Roberts, *New Perspectives on Faking in Personality Assessment* (págs. 151-173). New York, NY: Oxford University Press, Inc.
- Paulhus, D. L. y Harms, P. D. (2004). Measuring cognitive ability with the overclaiming technique. *Intelligence*, 32, 297-314. <https://doi.org/10.1016/j.intell.2004.02.001>
- Paulhus, D. L. y Holden, R. R. (2010). Measuring self-enhancement: from self-report to concrete behavior. En C. R. Agnew, D. E. Carlston, W. G. Graziano y J. R. Kelly (Eds.), *Then a Miracle Occurs: Focusing on Behavior in Social Psychological Theory and Research* (págs. 227-246). New York: Oxford University Press.
- Paulhus, D. L. y John, O. P. (1998). Egoistic and moralistic biases in self-perception: The interplay of self-deceptive styles with basic traits and motives. *Journal of Personality*, 66, 1025-1060. <https://doi.org/10.1111/1467-6494.00041>
- Paulhus, D. L. y Reid, D. B. (1991). Enhancement and denial in socially desirable responding. *Journal of Personality and Social Psychology*, 60, 307-317. <http://dx.doi.org/10.1037/0022-3514.60.2.307>

- Pauls, C. A. y Crost, N. W. (2005a). Cognitive ability and self-reported efficacy of self-presentation predict faking on personality measures. *Journal of Individual Differences*, 26, 194-206. <https://doi.org/10.1027/1614-0001.26.4.194>.
- Pauls, C. A. y Crost, N. W. (2005b). Effects of different instructional sets on the construct validity of the NEO-PI-R. *Personality and Individual Differences*, 39, 297-308. <https://doi.org/10.1016/j.paid.2005.01.003>
- \*Pavlov, G., Maydeu-Olivares, A. y Fairchild, A. J. (2019). Effects of applicant faking on forced-choice and Likert scores. *Organizational Research Methods*, 22, 710-739. <https://doi.org/10.1177/1094428117753683>
- Pennebaker, J. W., Francis, M. E. y Booth, R. J. (2001). *Linguistic Inquiry and Word Count (LIWC): LIWC2001*. Mahwah, NJ: Erlbaum.
- Peterson, M. H., Griffith, R. L. y Converse, P. D. (2009). Examining the role of applicant faking in hiring decisions: Percentage of fakers hired and hiring discrepancies in single-and multiple-predictor selection. *Journal of Business and Psychology*, 24, 373-386.
- Peterson, M. H., Griffith, R. L., Isaacson, J. A., O'Connell, M. S. y Mangos, P. M. (2011). Applicant faking, social desirability, and the prediction of counterproductive work behaviors. *Human Performance*, 24, 270-290. <https://doi.org/10.1080/08959285.2011.580808>
- Peterson, M. H., Griffith, R. L., O'Connell, M. S. y Isaacson, J. A. (2008). Examining faking in real job applicants: A within-subjects investigation of score changes across applicant and research settings. En R. L. Griffith y M. H. Peterson (Chairs), *Examining faking using within-subjects designs and applicant data*. Simposium realizado en la 23ª Annual Conference for the Society for Industrial and Organizational Psychology: San Francisco, CA.
- Pinder, C. C. (1998). *Motivation in Work Organizations*. Upper Saddle River, NJ: Prentice Hall.
- Poropat, A. E. (2009). A meta-analysis of the five-factor model of personality and academic performance. *Psychological Bulletin*, 135, 322-338.
- Postlethwaite, B. E. (2011). *Fluid ability, crystallized ability, and performance across multiple domains: a meta-analysis*. Tesis doctoral no publicada. University of Iowa, Iowa, IO.
- Raymark, P. H. y Tafero, T. L. (2009). Individual differences in the ability to fake on personality measures. *Human Performance*, 22, 86-103. <https://doi.org/10.1080/08959280802541039>

- Reeder, M. C. y Ryan, A. M. (2012). Methods for correcting for faking. En M. Ziegler, C. MacCann y R. D. Roberts (Eds.), *New Perspectives on Faking in Personality Assessment* (págs. 131-150). New York, NY: Oxford University Press.
- Rees, C. J. y Metcalfe, B. (2003). The faking of personality questionnaire results: who's kidding whom? *Journal of Managerial Psychology*, 18, 156-165.
- Robie, C., Brown, D. J. y Beaty, J. C. (2007). Do people fake on personality inventories? A verbal protocol analysis. *Journal of Business and Psychology*, 21, 489-509.
- Robie, C., Curtin, P. J., Foster, T. C., Phillips, H. L. IV., Zbylut, M. y Tetrick, L. E. (2000). The effect of coaching on the utility of response latencies in detecting fakers on a personality measure. *Canadian Journal of Behavioural Science*, 32, 226-233. <http://dx.doi.org/10.1037/h0087119>
- Robie, C., Tuzinski, K. A. y Bly, P. R. (2006). A survey of assessor beliefs and practices related to faking. *Journal of Managerial Psychology*, 21, 669-681. <https://doi.org/10.1108/02683940610690204>
- Robinson, M. D. y Ryff, C. D. (1999). The role of self-deception in perceptions of past, present, and future happiness. *Personality and Social Psychology Bulletin*, 25, 596-608. <https://doi.org/10.1177/0146167299025005005>
- Robson, S. M., Jones, A. y Abraham, J. (2008). Personality, faking, and convergent validity: a warning concerning warning statements. *Human Performance*, 21, 89-106. <https://doi.org/10.1080/08959280701522155>
- Roeder, S. y Paulhus, D. L. (2009, febrero). *Measuring consumer knowledge in the face of exaggeration and sabotage*. Póster presentado en el Meeting of the Society for Consumer Psychology, San Diego, CA.
- Rolland, J. P., y Mogenet, J. L. (2001). *Système de description en cinq dimensions (D5D). Manuel réservé aux psychologues*. París, Francia: Les Editions du Centre de Psychologie Appliquée.
- Rosse, J. G., Stecher, M. D., Miller, J. L. y Levin, R. A. (1998). The impact of response distortion on preemployment personality testing and hiring decisions. *Journal of Applied Psychology*, 83, 634-644. <http://dx.doi.org/10.1037/0021-9010.83.4.634>
- Roth, P. L., BeVier, C. A., Switzer III, F. S. y Schippmann, J. S. (1996). Meta-analyzing the relationship between grades and job performance. *Journal of Applied Psychology*, 81, 548-556. <http://dx.doi.org/10.1037/0021-9010.81.5.548>
- Rothstein, M. G. y Goffin, R. D. (2006). The use of personality measures in personnel selection: What does current research support? *Human Resource Management Review*, 16, 155-180. <https://doi.org/10.1016/j.hrmr.2006.03.004>

- Ruch, F. L. (1942). A technique for detecting attempts to fake performance on a selfinventory type of personality test. En Q. McNemar y M. A. Merrill (Eds.), *Studies in Personality* (págs. 229-234). New York, NY: McGraw-Hill.
- \*Rusmore, J. T. (1956). Fakability of the Gordon Personal Profile. *Journal of Applied Psychology*, 40, 175-177. <http://dx.doi.org/10.1037/h0042524>
- Sackeim, H. A. y Gur, R. C. (1978). Self-deception, self-confrontation, and consciousness. En Schwartz, G. E. y Shapiro, D. (Eds.), *Consciousness and self-regulation* (págs. 139-197). Boston, MA: Springer.
- Sackeim, H. A. y Gur, R. C. (1979). Self-deception, other-deception, and self-reported psychopathology. *Journal of Consulting and Clinical Psychology*, 47, 213-215. <http://dx.doi.org/10.1037/0022-006X.47.1.213>
- Sackett, P. R., Burris, L. R. y Callahan, C. (1989). Integrity testing for personnel selection: An update. *Personnel Psychology*, 42, 491-529. <https://doi.org/10.1111/j.1744-6570.1989.tb00666.x>
- Salgado, J. F. (1997a). The five factor model of personality and job performance in the European Community. *Journal of Applied Psychology*, 82, 30-43. <http://dx.doi.org/10.1037/0021-9010.82.1.30>
- Salgado, J. V. (1997b). VALCOR: A program for estimating standard error, confidence intervals and probability of corrected validity. *Behavior Research, Methods, Instruments, & Computers*, 29, 464-467. <https://doi.org/10.3758/BF03200601>
- Salgado, J. F. (1998). *Manual técnico del inventario de personalidad de cinco factores (IP/5F)*. Santiago de Compostela: Tórculo
- Salgado, J. F. (2003). Predicting job performance using FFM and non-FFM personality measures. *Journal of Occupational and Organizational Psychology*, 76, 323-346. <https://doi.org/10.1348/096317903769647201>
- Salgado, J. F. (2005). Personalidad y deseabilidad social en contextos organizacionales: implicaciones para la práctica de la psicología del trabajo y las organizaciones. *Papeles del Psicólogo*, 26, 115-128.
- Salgado, J. F. (2010). *Escalas de desempeño académico: CDTE, CDCE y CDAN*. Manuscrito no publicado, Departamento de Psicología Social, Básica y Metodología, Universidad de Santiago de Compostela, Santiago de Compostela, España.
- Salgado, J. F. (2014). *Reliability, construct and criterion validity of the Quasipsative Personality Inventory (QI5F/Tri)*. Manuscrito no publicado, Departamento de



Psicología Social, Básica y Metodología, Universidad de Santiago de Compostela, Santiago de Compostela, España.

- Salgado, J. F. (2015). Estimating Coefficients of Equivalence and Stability for Job Performance Ratings: The importance of controlling for transient error on criterion measurement. *International Journal of Selection and Assessment*, 23, 37-44. doi: 10.1111/ijsa.12093
- Salgado, J. F. (2016). A theoretical model of psychometric effects of faking on assessment procedures: Empirical findings and implications for personality at work. *International Journal of Selection and Assessment*, 24, 209-228. <https://doi.org/10.1111/ijsa.12142>
- Salgado, J. F. (2017). Moderator effects of job complexity on the validity of forced-choice personality inventories for predicting job performance. *Revista de Psicología del Trabajo y de las Organizaciones*, 33, 229-238.
- Salgado, J. F., Anderson, N. y Táuriz, G. (2015). The validity of ipsative and quasi-ipsative forced-choice personality inventories for different occupational groups: A comprehensive meta-analysis. *Journal of Occupational and Organizational Psychology*, 88, 797-834. <https://doi.org/10.1111/joop.12098>
- Salgado, J. F. y De Fruyt, F. (2005). Personality in personnel selection. En A. Evers, N. Anderson y O. Voskuijl (Eds.), *The Blackwell Handbook of Personnel Selection* (págs. 174-198). Oxford, UK: Blackwell.
- \*Salgado, J. F. y Lado, M. (2018). Faking resistance of a quasi-ipsative forced-choice personality inventory without algebraic dependence. *Revista de Psicología del Trabajo y de las Organizaciones*, 34, 213-216. <http://dx.doi.org/10.5093/jwop2018a23>
- Salgado, J. F. y Moscoso, S. (2008). Selección de personal en la empresa y las AAPP: de la visión tradicional a la visión estratégica. *Papeles del Psicólogo*, 29, 16-24.
- Salgado, J. F. y Moscoso, S. (2018). *Selección, evaluación y desarrollo de personal: fundamentos*. Manuscrito no publicado, Departamento de Ciencia Política y Sociología, Universidad de Santiago de Compostela, Santiago de Compostela, España.
- Salgado, J. F. y Moscoso, S. (2019). The validity of general mental ability for five performance criteria: Hunter and Hunter (1984) revisited. *Frontiers in Psychology*, under review.
- Salgado, J. F. y Táuriz, G. (2014). The Five-Factor Model, forced-choice personality inventories and performance: A comprehensive meta-analysis of academic and occupational validity studies. *European Journal of Work and Organizational Psychology*, 23, 3-30. <https://doi.org/10.1080/1359432X.2012.716198>

- Saville, P., Holdsworth, R., Nyfield, G., Cramp, L. y Mabey, W. (1984). *The Occupational Personality Questionnaires (OPQ)*. London: Saville & Holdsworth Ltd.
- Saville, P. y Willson, E. (1991). The reliability and validity of normative and ipsative approaches in the measurement of personality. *Journal of Occupational Psychology*, 64, 219-238. <https://doi.org/10.1111/j.2044-8325.1991.tb00556.x>
- Schmidt, F. L. y Hunter, J. E. (2015). *Methods of Meta-Analysis: Correcting Error and Bias in Research Findings*. Thousand Oaks, CA: Sage.
- Schmidt, F. L. y Le, H. (2004). *Software for the Hunter-Schmidt meta-analysis methods*. Iowa City, IA: Department of Management and Organizations, University of Iowa.
- Schmidt, F. L., Le, H. y Ilies, R. (2003). Beyond alpha: An empirical examination of the effects of different sources of measurement error on reliability estimates for measures of individual-differences constructs. *Psychological Methods*, 8, 206-224. <http://dx.doi.org/10.1037/1082-989X.8.2.206>
- Schmit, M. J. y Ryan, A. M. (1993). The Big Five in personnel selection: Factor structure in applicant and nonapplicant populations. *Journal of Applied Psychology*, 78, 966-974. <http://dx.doi.org/10.1037/0021-9010.78.6.966>
- Schmitt, N. y Klimoski, R. (1991). *Research Methods in Human Resources Management*. Cincinnati, OH: South-Western Publishing Co.
- Schmitt, N. y Oswald, F. L. (2006). The impact of corrections for faking on the validity of noncognitive measures in selection settings. *Journal of Applied Psychology*, 91, 613-621. <http://dx.doi.org/10.1037/0021-9010.91.3.613>
- Schuessler, K., Hittle, D. y Cardascia, J. (1978). Measuring responding desirably with attitude-opinion items. *Social Psychology*, 224-235. <http://dx.doi.org/10.2307/3033559>
- \*Schwab, D. P. (1971). Issues in response distortion studies of personality inventories: A critique and replicated study. *Personnel Psychology*, 24, 637-647. <http://dx.doi.org/10.1111/j.1744-6570.1971.tb00377.x>
- Seisdedos, N. (1993). Personnel selection, questionnaires, and motivational distortion: An intelligent attitude of adaptation. En H. Schuler, J. L. Farr y M. Smith (Eds.), *Personnel selection and assessment* (págs. 91-108). Hillsdale, NJ: Erlbaum.
- \*Sheppard, C., Bates, C., Fracchia, J. y Merlis, S. (1974). Psychological need structures of law enforcement officers. *Psychological Reports*, 35, 583-586. <https://doi.org/10.2466/pr0.1974.35.1.583>



- Shiflett, S. y Cohen, S. L. (1982). The shifting salience of valence and instrumentality in the prediction of perceived effort, satisfaction and turnover. *Motivation and Emotion*, 6, 65-77.
- SHL Group. (1993b). OPQ concept model: Manual and user's guide. Thames Ditton, United Kingdom: Author
- Shore, L. M. y Wayne, S. J. (1993). Commitment and employee behavior: Comparison of affective commitment and continuance commitment with perceived organizational support. *Journal of Applied Psychology*, 78, 774-780. <http://dx.doi.org/10.1037/0021-9010.78.5.774>
- Society for Industrial and Organizational Psychology (2003). *Principles for the Validation and Use of Personnel Selection Techniques*. Bowling Green: APA.
- Smith, G. M. (1967). Usefulness of peer ratings of personality in educational research. *Educational and Psychological Measurement*, 27, 967-984. <https://doi.org/10.1177/001316446702700445>
- Smith, D. B. y Ellingson, J. E. (2002). Substance versus style: A new look at social desirability in motivating contexts. *Journal of Applied Psychology*, 87, 211-219. <http://dx.doi.org/10.1037/0021-9010.87.2.211>
- Smith, D. B., Hanges, P. J. y Dickson, M. W. (2001). Personnel selection and the five-factor model: Reexamining the effects of applicant's frame of reference. *Journal of Applied Psychology*, 86, 304-315. <http://dx.doi.org/10.1037/0021-9010.86.2.304>
- Smith, D. B. y McDaniel, M. (2012). Questioning old assumptions: Faking and the personality-performance relationship. En M. Ziegler, C. MacCann y R. D. Roberts (Eds.), *New perspectives on faking in personality assessment* (págs. 53-69). Oxford: Oxford University Press.
- Snell, A. F., Sydell, E. J. y Lueke, S. B. (1999). Towards a theory of applicant faking: Integrating studies of deception. *Human Resource Management Review*, 9, 219-242. [https://doi.org/10.1016/S1053-4822\(99\)00019-4](https://doi.org/10.1016/S1053-4822(99)00019-4)
- \*Spector, A. J. (1957). The Attitudes Test in Human Relations (ATHURE). *Journal of Applied Psychology*, 41, 209-213. <http://dx.doi.org/10.1037/h0048266>
- Stark, S., Chernyshenko, O. S., Chan, K. Y., Lee, W. C. y Drasgow, F. (2001). Effects of the testing situation on item responding: Cause for concern. *Journal of Applied Psychology*, 86, 943-953.
- \*Stollak, G. E. (1965). EPPS performance under social desirability instructions: college females. *Psychological Reports*, 16, 119-122. <http://dx.doi.org/10.1037/h0022382>

- Swanson, A. y Ones, D. S. (2002). *The effects of individual differences and situational variables on faking*. Póster presentado en la Annual Conference of the Society for Industrial and Organizational Psychology, Toronto, Canada.
- Táuriz, G. (2011). Falseamiento y validez de las medidas de personalidad en contextos académicos. *Revista de Psicología del Trabajo y de las Organizaciones*, 27, 103-115.
- Tenopyr, M. L. (1988). Artifactual reliability of forced-choice scales. *Journal of Applied Psychology*, 73, 749-751. <http://dx.doi.org/10.1037/0021-9010.73.4.749>
- Tett, R.P., Anderson, M.G., Ho, C.L., Yang, T.S., Huang, L. y Hanvongse, A. (2006). Seven nested questions about faking on personality tests: An overview and interactionist model of item-level response distortion. En R. L. Griffith y M. H. Peterson (Eds.), *A closer examination of applicant faking behavior* (págs. 43-83). Greenwich, CT: Information Age Publishing, Inc.
- Tett, R. P., Christiansen, N. D., Robie, C. y Simonet, D. V. (2011, mayo). *International survey of personality test use: An American baseline*. Comunicación presentada en la 15ª Conference of the European Association of Work and Organizational Psychology, Maastricht, Holanda.
- Tett, R. P. y Simonet, D. V. (2011). Faking in personality assessment: A “multisaturation” perspective on faking as performance. *Human Performance*, 24, 302-321. <https://doi.org/10.1080/08959285.2011.597472>
- Thomson, G. H. (1938). The influence of univariate selection on the factorial analysis of ability. *British Journal of Psychology*, 28, 451-459.
- Thomson, G. H. (1939). *The factorial analysis of human ability*. New York, NY: Houghton Mifflin.
- Thomson, G. H. (1944). The applicability of Karl Pearson’s selection formulae in follow-up experiments. *British Journal of Psychology: General Section*, 34, 105.
- Tibbetts, S. G. (1997). Shame and rational choice in offending decisions. *Criminal Justice and Behavior*, 24, 234-255. <https://doi.org/10.1177/0093854897024002006>
- Travers, R. M. (1951). A critical review of the validity and rationale of the forced-choice technique. *Psychological Bulletin*, 48, 62-70. <http://dx.doi.org/10.1037/h0055263>
- Treviño, L. K. y Youngblood, S. A. (1990). Bad apples in bad barrels: A causal analysis of ethical decision-making behavior. *Journal of Applied Psychology*, 75, 378-385. <http://dx.doi.org/10.1037/0021-9010.75.4.378>
- Tupes, E., y Christal, R. (1961). Recurrent personality factors based on trait ratings. *Journal of Personality*, 60, 225-251.

- Turnley, W. H. y Bolino, M. C. (2001). Achieving desired images while avoiding undesired images: exploring the role of self-monitoring in impression management. *Journal of Applied Psychology*, 86, 351-360. <http://dx.doi.org/10.1037/0021-9010.86.2.351>
- Tversky, A. y Kahneman, D. (1986). Judgment under uncertainty: Heuristics and biases. *Judgment and Decision Making: An Interdisciplinary Reader*, 38-55.
- \*Underhill, C. M., Bearden, R. M. y Chen, H. T. (2008). *Evaluation of the Fake Resistance of a Forced-choice Paired-comparison Computer Adaptive Personality Measure* (N° NPRST-TR-08-2). Navy Personnel Research Studies and Technology. Millington, Tennessee, TN.
- Van Hooft, E. A. y Born, M. P. (2012). Intentional response distortion on personality tests: Using eye-tracking to understand response processes when faking. *Journal of Applied Psychology*, 97, 301-316. <http://dx.doi.org/10.1037/a0025711>
- Van Iddekinge, C. H., Raymark, P. H. y Roth, P. L. (2005). Assessing personality with a structured employment interview: Construct-related validity and susceptibility to response inflation. *Journal of Applied Psychology*, 90, 536-552. <http://dx.doi.org/10.1037/0021-9010.90.3.536>
- VandeWalle, D. (1997). Development and validation of a work domain goal orientation instrument. *Educational and Psychological Measurement*, 57, 995-1015. <https://doi.org/10.1177/0013164497057006009>
- \*Vasilopoulos, N. L., Cucina, J. M., Dyomina, N. V., Morewitz, C. L. y Reilly, R. R. (2006). Forced-choice personality tests: A measure of personality and cognitive ability? *Human Performance*, 19, 175-199. [https://doi.org/10.1207/s15327043hup1903\\_1](https://doi.org/10.1207/s15327043hup1903_1)
- Ventura, M. (2012). The detection of faking through word use. En M. Ziegler, C. MacCann y R. D. Roberts (Eds.) *New Perspectives on Faking in Personality Assessment* (págs. 165-173). New York, NY: Oxford University Press.
- Vernon, P. E. (1934). The attitude of the subject in personality testing. *Journal of Applied Psychology*, 18, 165-177. <http://dx.doi.org/10.1037/h0074033>
- Viswesvaran, C. y Ones, D. S. (1999). Meta-analyses of fakability estimates: Implications for personality measurement. *Educational and Psychological Measurement*, 59, 197-210. <https://doi.org/10.1177/00131649921969802>
- Vroom, V. H. (1964). *Work and Motivation*. New York, NY: Wiley.
- Walczyk, J. J., Roper, K. S., Seemann, E. y Humphrey, A. M. (2003). Cognitive mechanisms underlying lying to questions: Response time as a cue to deception. *Applied Cognitive Psychology: The Official Journal of the Society for*

- Applied Research in Memory and Cognition*, 17, 755-774.  
<https://doi.org/10.1002/acp.914>
- Walczyk, J. J., Schwartz, J. P., Clifton, R., Adams, B., Wei, M. I. N. y Zha, P. (2005). Lying person-to-person about life events: A cognitive framework for lie detection. *Personnel Psychology*, 58, 141-170. <https://doi.org/10.1111/j.1744-6570.2005.00484.x>
- Wiggins, J. S. (1964). Convergences among stylistic response measures from objective personality tests. *Educational and Psychological Measurement*, 24, 551-562. <https://doi.org/10.1177/001316446402400310>
- Wrensen, L. B. y Biderman, M. D. (2005, abril). *Factors related to faking ability: A structural equation model application*. Comunicación presentada en la 20ª Annual Conference of the Society for Industrial and Organizational Psychology, Los Angeles, CA.
- Zavala, A. (1965). Development of the forced-choice rating scale technique. *Psychological Bulletin*, 63, 117-124.  
<http://dx.doi.org/10.1037/h0021567>
- Zickar, M. J. (2000). Modeling faking on personality tests. En D. R. Ilgen y C. L. Hulin (Eds.), *Computational modeling of behavior in organizations: The third scientific discipline* (págs. 95-113). Washington, DC: American Psychological Association.  
<http://dx.doi.org/10.1037/10375-005>
- Zickar, M. J. y Gibby, R. E. (2006). A history of faking and socially desirable responding on personality tests. En R. L. Griffith y M. H. Peterson. *A Closer Examination of Applicant Faking Behavior* (págs. 21-42). Greenwich, CT: Information Age Publishing, Inc.
- Zickar, M. J., Gibby, R. E. y Robie, C. (2004). Uncovering faking samples in applicant, incumbent, and experimental data sets: An application of mixed-model item response theory. *Organizational Research Methods*, 7, 168-190.  
<https://doi.org/10.1177/1094428104263674>
- Zickar, M. J., Rosse, J. y Levin, R. (1996). Modeling of faking in a selection context. En C. L. Hulin (Chair), *The third scientific discipline: Computational modeling in organizational research*. Simposium realizado en el Annual meeting of the Society for Industrial and Organizational Psychology, San Diego, CA.
- Zickar, M., Rosse, J., Levin, R. y Hulin, C. (1996, abril). *Modeling the effects of faking on personality scales*. Comunicación presentada en la 11ª Annual Conference of the Society for Industrial and Organizational Psychology, San Diego, CA.
- Zickar, M. J. y Sliter, K. A. (2012). Searching for unicorns. Item response theory-based solutions to the faking problem. En M. Ziegler, C. Maccann y R. D. Roberts

(Eds.), *New Perspectives on Faking in Personality Assessment* (págs. 113-130). New York, NY: Oxford University Press.

Ziegler, M. y Buehner, M. (2009). Modeling socially desirable responding and its effects. *Educational and Psychological Measurement*, 69, 548-565. <https://doi.org/10.1177/0013164408324469>

Ziegler, M., Maaß, U., Griffith, R. y Gammon, A. (2015). What is the nature of faking? Modeling distinct response patterns and quantitative differences in faking at the same time. *Organizational Research Methods*, 18, 679-703. <https://doi.org/10.1177/1094428115574518>

Ziegler, M., MacCann, C. y Roberts, R. D. (2012). Faking: Knowns, unknowns, and points of contention. En M. Ziegler, C. MacCann y R. D. Roberts (Eds.), *New perspectives on faking in personality assessment* (págs. 3-16). New York, NY: Oxford University Press.

Ziegler, M., Toomela, A. y Buehner, M. (2009). A reanalysis of Toomela (2003): Spurious measurement error as cause for common variance between personality factors. *Psychological Science Quarterly*, 51, 61-75.

Ziegler, M., MacCann, C. M. y Roberts, R. D. (2012). *New Perspectives on Faking in Personality Assessment*. New York, NY: Oxford University Press.